

Reassessing the Boundary between Classical and Nonclassical for Individual Quantum Processes

Yujie Zhang^{1,2,3,*} David Schmid^{2,†} Yilè Yīng^{2,3,‡} and Robert W. Spekkens²

¹*Institute for Quantum Computing, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*

²*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, Ontario N2L 2Y5, Canada*

³*Department of Physics and Astronomy, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada*



(Received 18 April 2025; revised 28 January 2026; accepted 14 April 2026; published 4 June 2026)

There is a received wisdom about where to draw the boundary between classical and nonclassical for various types of quantum processes. For multipartite states, it is the divide between separable and entangled; for channels, the divide between entanglement breaking and not; for sets of measurements, the divide between compatible and incompatible; for assemblages, the divide between unsteerable and steerable. However, these choices have not been motivated by any unified notion of what it means to be classically explainable. One well-motivated notion of classical explainability is the one based on generalized noncontextuality: A set of circuits is classically explainable if a generalized-noncontextual ontological model can realize the statistics they generate. In this work, we show that this notion can be leveraged to define a classical-nonclassical divide for individual quantum processes of arbitrary type. A set of measurements, for example, is judged to be classical if and only if a particular set of circuits—the one obtained by contracting these measurements with every possible quantum state—is classically explainable in the sense just articulated. We begin the task of characterizing where the classical-nonclassical divide lies according to this proposal for a variety of different types of processes. In particular, we show that all of the following are judged to be nonclassical: every entangled state, every set of incompatible measurements, every non-entanglement-breaking channel, and every steerable assemblage. Our proposal differs from the received wisdom, however, insofar as it also judges certain subsets of the complementary classes to be nonclassical, including certain separable states, compatible sets of measurements, entanglement-breaking channels, and unsteerable assemblages. Finally, we prove structure theorems characterizing the classical-nonclassical divide based on whether a process admits of a specific type of frame representation.

DOI: [10.1103/vqfz-wzjg](https://doi.org/10.1103/vqfz-wzjg)

Subject Areas: Quantum Physics,
Quantum Information

I. INTRODUCTION

A recurring question of interest in quantum information and foundations is to determine which processes (states, measurements, sets of measurements, channels, and so on) can be reasonably said to be classical, and which must be acknowledged to be nonclassical (that is, genuinely quantum).

In much of the literature, proposals for the placement of the classical-nonclassical divide have been introduced on a case-by-case basis for different types of processes.

For instance, it is commonly taken that for bipartite states, it is the divide between separable and entangled, for channels, it is the divide between entanglement breaking and non-entanglement-breaking, for multi-measurements, it is the divide between the set of measurements being compatible and being incompatible, and so on.

These proposals are often based on intuitive notions of what should count as a classical kind of correlation or process. However, the question of whether all such proposals are consistent with a single principle has remained largely unexplored [1]. Indeed, the resulting conceptions of nonclassicality can be distinct for different types of processes, and we will even show that they are sometimes in conflict with one another.

Seeking to define classicality across different types of processes using a unified notion of classical explainability is especially important, given that many widely held intuitions about which features of quantum theory count as evidence of genuine quantumness are inconsistent with well-motivated notions of classical explainability.

*Contact author: yujie4physics@gmail.com

†Contact author: dschmid1@pitp.ca

‡Contact author: yying@pitp.ca

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

For example, there exist classical theories [5–7] that satisfy various principles, such as locality and generalized non-contextuality, and that nonetheless contain analogs of entanglement, non-entanglement-breaking channels, and incompatible measurements, not to mention steering, interference, no cloning, and a wide variety of other phenomena that are commonly thought to be challenging to explain classically.

The question of where the classical-nonclassical divide lies is sometimes mistaken for one that is merely about semantics, that is, merely about how we use the term “classical.” This is not the case because the answer to this question has a significant impact on how one pursues foundational and applied research in quantum theory. For instance, the answer impacts how one pursues the project of identifying quantum-over-classical advantages across various tasks (computational, cryptographic, metrological, thermodynamic, etc.), because one expects to find such advantages only in protocols that leverage nonclassical resources. In other words, a desideratum for a good proposal on where to place the classical-nonclassical divide is that any advantage should require resources that are deemed nonclassical by the proposal. Similarly, the placement of the classical-nonclassical divide is important for identifying the manner in which the quantum worldview differs from the classical worldview and hence for correctly identifying the true innovation of quantum theory over its classical predecessors.

Here, we introduce an approach that defines the classical-nonclassical divide for a process based on a single principle, namely, the idea that classical explainability of operational statistics corresponds to realizability by a generalized-noncontextual ontological model [8,9].

Since our framework is based on the notion of a generalized-noncontextual ontological model, it is worth reiterating the motivations for why realizability in such a model constitutes a good notion of classical explainability for operational statistics. The ontological models framework captures the conventional manner of providing a realist account of an operational theory [10]. Demanding that a realist account should satisfy the principle of generalized noncontextuality can be motivated by a methodological principle for theory construction due to Leibniz. This principle has proven its utility in physics through Einstein’s use of it in his development of the special and general theories of relativity [11]. Finally, we note that a generalized-noncontextual ontological representation is equivalent to a positive quasiprobability representation [9,12,13] and coincides with the natural notion of classical explainability in the framework of generalized probabilistic theories (GPTs [14,15]), namely, linear embeddability into the strictly classical (simplicial) GPT [9,13].

This notion of classical explainability also has pragmatic advantages over some of its competitors. For one, it is noise robust and is therefore testable in real experiments where

the ideal of noiselessness is never realized [16]. This is in contrast to the Kochen-Specker notion of noncontextuality [17], as argued in Refs. [16,18–20].

It also has universal applicability: Given a circuit with arbitrary causal structure, one can ask whether there exists a generalized-noncontextual ontological model that can reproduce the statistics generated by that circuit. In contrast, standard Bell inequalities [21] apply only to circuits having one specific causal structure. Because of its universal applicability, one can ask: For any operational phenomenon, precisely which aspects of it fail to admit of a generalized-noncontextual ontological model, and hence, which aspects resist classical explanation? Results establishing facts about the boundary between what is classically explainable and what is not are now known for the phenomenology of quantum computation [22,23], state discrimination [24–27], interference [7,28–30], compatibility [31–33], uncertainty relations [34], metrology [35], thermodynamics [35–37], weak values [38,39], coherence [40–42], quantum Darwinism [43], information processing and communication [44–50], cloning [51], broadcasting [52], pre- and postselection paradoxes [53], randomness certification [54], psi-epistemicity [55], macroscopic realism [56], and Bell [10,57] and Kochen-Specker scenarios [19,58–63]. Aiding this research is a growing set of powerful techniques for certifying the failure to admit of a noncontextual ontological model [9,13,16,64–71].

However, in all work to date, the existence of a generalized-noncontextual ontological model provides a notion of classical explainability for a closed circuit, since only such a circuit can generate operational statistics, and it is on the basis of these statistics that classical explainability is assessed. In contrast, no notion of classicality has been given for a single process within a circuit, such as a single quantum state, measurement, or channel. In this work, we propose a way of doing so.

The intuitive idea is straightforward: A given quantum process is nonclassical if and only if there exists a quantum circuit that leverages it nontrivially to generate statistics that are inconsistent with a generalized-noncontextual ontological model. In Sec. III, we formalize this definition. In particular, we explain what it means to “leverage in a nontrivial way” the process, and then we simplify the definition considerably, by showing that rather than considering all quantum circuits containing the process, one needs only to consider circuits where product states are prepared on its inputs and product measurements are implemented on its outputs.

We then undertake the task of characterizing which quantum preparations, measurements, and transformations are nonclassical, and which are classical. (This paper focuses only on quantum theory, although our approach could be applied to an arbitrary generalized probabilistic theory.)

We prove structure theorems characterizing the structure of any classical process in terms of the existence of a

particular kind of product dual-frame representation. We then use this to prove a number of interesting results, including the following:

- (1) A single state on a unipartite system is always classical.
- (2) Any multi-state consisting of a set of distinct pure states on a unipartite system is nonclassical if and only if the states are linearly dependent in the operator space (and therefore, any minimal tomographically complete set of pure states is classical).
- (3) Any rank-1 positive-operator-valued measure (POVM) is nonclassical if and only if its POVM elements are linearly dependent in the operator space [and therefore, any symmetric informationally complete (SIC) POVM is classical].
- (4) The set of classical multi-states (respectively, measurements) is nonconvex.
- (5) Every steerable assemblage is nonclassical, as are certain unsteerable assemblages.
- (6) Every multi-measurement consisting of a set of incompatible measurements is nonclassical, as are certain sets of compatible measurements (including certain singleton sets, i.e., single measurements).
- (7) Every entangled state is nonclassical, as are certain separable states.
- (8) For two-qubit systems, a state is nonclassical if and only if it is entangled.
- (9) All non-entanglement-breaking channels are nonclassical, as are certain entanglement-breaking channels.

Note that “multi-state” and “multi-measurement” are terms we use in this paper to denote a type of single process associated with a set of states or a set of measurements, respectively (see Sec. II A).

In addition, we show that the task of deciding nonclassicality for a quantum process of arbitrary type (including channels and instruments) can be reduced to the corresponding task for an associated multipartite state.

Together, these results provide a useful characterization of the nonclassicality of arbitrary processes. The essential results of the paper are summarized in Fig. 1 below.

In a companion paper [72], we provide further tools based on semidefinite programming that help one to witness and quantify the nonclassicality of single states and measurements. We do so both in a theory-dependent way (analogous to entanglement witnesses) and in a theory-independent way (using noncontextuality inequalities).

In another companion paper [73], we focus on bipartite states and investigate how our notion of nonclassicality relates to entanglement, steering, and nonlocality. We provide tools to certify these quantum resources via a hierarchy of noncontextuality inequalities. Moreover, we present an experimental demonstration of this novel entanglement-certification approach and contrast it with existing methods.

We end with two points about terminology.

First, we wish to head off a possible point of confusion: The term “quantum” is sometimes taken as a synonym for nonclassical, and other times taken as an enveloping term

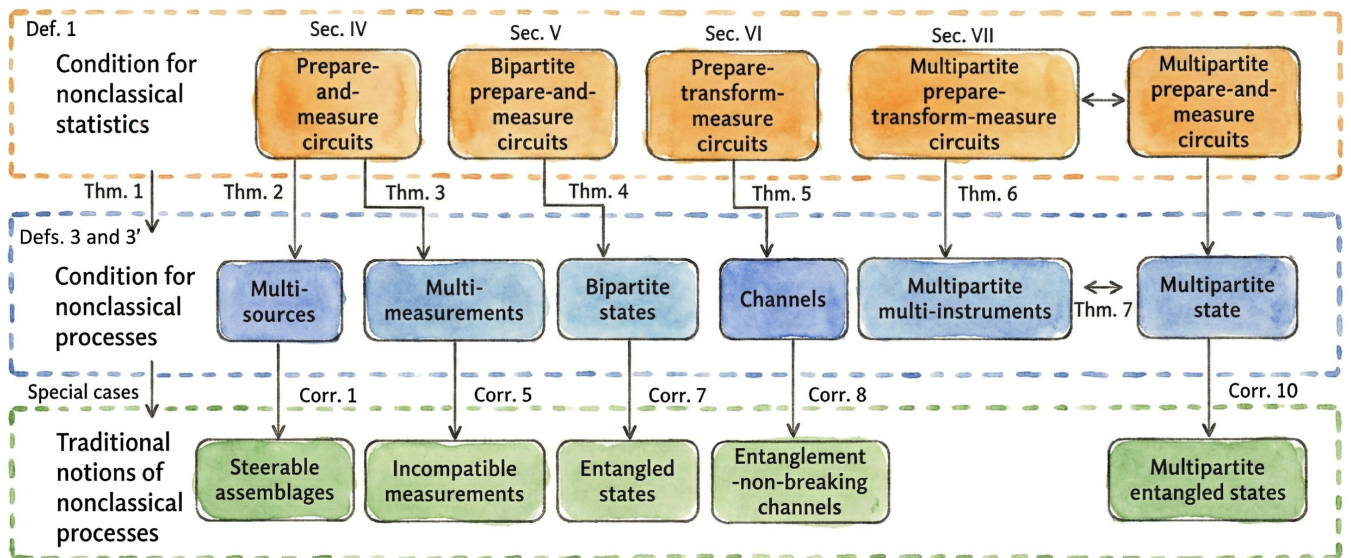


FIG. 1. Schematic showing the main results of the paper. Yellow: in each section, we study the class of quantum circuits built from the type of target process of interest together with dual processes (Definition 2) that close the circuit; the condition for the resulting statistics to be classically explainable is realizability in a noncontextual ontological model (Definition 1). Blue: a process is then defined to be classical if the resulting statistics from the set of circuits obtained by contracting it with all possible dual processes are classically explainable (Definitions 3 and 3’); the classical-nonclassical divide for the processes of each type is then characterized by the corresponding structure theorems. Green: traditional notions of nonclassicality are recovered as special cases, i.e., as sufficient conditions for the proposed notion of nonclassicality.

that includes classical and nonclassical as special cases [74]. In this work, we use the term quantum in the more inclusive sense. This is because we are studying the question of which subsets of phenomena described by quantum theory are classically explainable and which are not. We use the term “nonclassical” or “intrinsically quantum” to signal the impossibility of classical explanation [75].

Second, the term “classical” is quite overloaded in the literature. Although we shall continue to use the term throughout this article (where its meaning should be clear from context and from our explicit definitions), we suggest the term Leibniz classical be used by anyone who seeks a more specific term (e.g., if referring to our notion in a broader setting where it might not be clear from context to what the word classical refers).

II. PRELIMINARIES

A. Types of quantum processes

A generic process in quantum theory is simply a transformation with various input and output systems, which may be quantum or classical. A classical system that is an input is generally called a setting variable and one that is an output is generally called an outcome variable. The most common examples of processes in quantum theory are states, effects, measurements, channels, and instruments. However, it is often useful to consider more general processes, including those listed in the Table. I.

Most of these processes will be familiar to the quantum information theorist. However, the terminology of sources, multi-sources, multi-measurements, and multi-channels is

TABLE I. Examples of various types of processes, their representations in quantum theory, and their representations as substochastic matrices in an ontological model. Throughout this paper, we denote classical outcome variables by a, b, c and setting variables by x, y, z , while quantum inputs and quantum outputs are denoted by the capital letters A, B, C .

Process type	Quantum rep	Ontological rep
State	ρ^A	$p(\lambda_A \rho^A)$
Effect	M^A	$p(M^A \lambda_A)$
Channel	$\mathcal{E}^{B A}$	$p(\lambda_B \lambda_A)$
Source	$\{p(a)\rho_a^A\}_a$	$\{p(a\lambda_A)\}_a$
Measurement	$\{M_b^A\}_b$	$\{p(b \lambda_A)\}_b$
Instrument	$\{\mathcal{F}_c^{B A}\}_c$	$\{p(c\lambda_B \lambda_A)\}_c$
Multi-state	$\{\rho_{\cdot x}^A\}_x$	$\{p(\lambda_A x)\}_x$
Multi-channel	$\{\mathcal{E}_{\cdot z}^{B A}\}_z$	$\{p(\lambda_B z\lambda_A)\}_z$
Multi-source	$\{p(a x)\rho_{a x}^A\}_{a,x}$	$\{p(a\lambda_A x)\}_{a,x}$
Multi-measurement	$\{M_{b y}^A\}_{b,y}$	$\{p(b y\lambda_A)\}_{b,y}$
Multi-instrument	$\{\mathcal{F}_{c z}^{B A}\}_{c,z}$	$\{p(c\lambda_B z\lambda_A)\}_{c,z}$
Multipartite state	$\rho^{A_1,\dots,A_n}$	$p(\lambda_1, \dots, \lambda_n)$

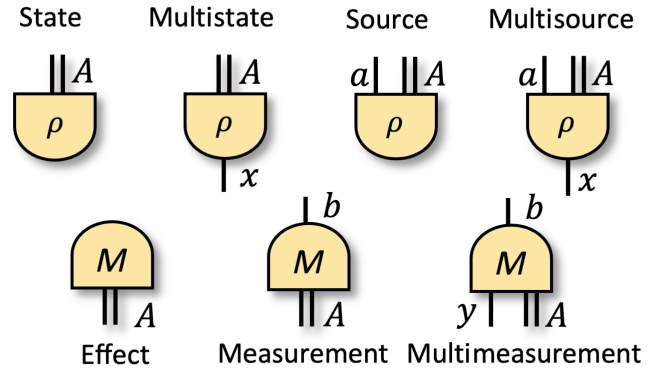


FIG. 2. Examples of processes with a single quantum output A (top) and with a single quantum input A (bottom) on a single quantum system. Throughout this paper, quantum systems are represented by double wires, whereas classical systems are represented by single wires.

relatively new, so we explain these terms below. These processes are easily understood via a diagrammatic representation, as in Fig. 2.

The type of a process is simply a specification of its input and output systems, as well as a specification of whether each of these is classical or quantum. Diagrammatically, we will depict quantum systems by double wires and classical systems by single wires (see, e.g., Fig. 2). In this approach, each individual input and output of a process (depicted by a wire in the diagram) is considered a fundamental system, having no subsystem structure, so that a composite system is always described by a collection of inputs or a collection of outputs.

Each quantum input and output for a given process is associated with a Hilbert space. For a quantum system A , we denote this Hilbert space by $\mathcal{H}^A$ with superscript A . We will denote the set of all linear operators on $\mathcal{H}$ as $\mathcal{L}(\mathcal{H})$, the set of all Hermitian operators on $\mathcal{H}$ as $\text{Herm}(\mathcal{H})$, the set of all density operators [i.e., all $\rho \in \text{Herm}(\mathcal{H})$ such that $\rho \geq 0$ and $\text{Tr}[\rho] = 1$] as $\mathcal{S}(\mathcal{H})$, and the set of all effects [i.e., all $M \in \text{Herm}(\mathcal{H})$ such that $0 \leq M \leq 1$] as $\mathcal{M}(\mathcal{H})$. Formally, an effect on system A is a map $\mathcal{M}^A(X^A) := \text{Tr}[M^A X^A]$, i.e., a channel with no output. Nonetheless, we will often simply refer to the operator M^A itself as an effect.

It is useful to classify the types of processes by their quantum inputs and outputs, such that the only differences among the processes in a given class are in their classical outcome and setting variables. For instance, processes with a single quantum output and no quantum input form a class. We will refer to these as processes of the preparation variety. In a similar fashion, the class of processes with a single quantum input and no quantum output will be termed processes of the measurement variety, and the class of processes with both a quantum input and a quantum output will be termed processes of the transformation variety. We will sometimes qualify the terms describing these varieties to stipulate the subsystem structure of the quantum inputs

or the quantum outputs; for example, we will speak of processes of the bipartite preparation variety.

We begin with processes of the preparation variety, i.e., those having a single quantum output A . Four types of processes in this class are depicted in the top row of Fig. 2. A single deterministic preparation on a quantum system A is represented by a density operator (or state) $\rho \in \mathcal{S}(\mathcal{H}^A)$. A process with a classical input x (a setting variable) that determines which of a set of preparations is implemented on A can be represented by a set of states $\{\rho_{a|x}^A\}_x$. We refer to such a process as a multi-state process. Next, consider a process in which a variable a is sampled at random from a distribution $p(a)$ and then the preparation represented by the state ρ_a^A is implemented on A , and the variable a becomes a classical output of the process (an outcome variable) that flags which state was prepared. Such a process is represented by a set of unnormalized states $\{p(a)\rho_a^A\}_a$, and we refer to it as a source (it is sometimes termed an ensemble). Finally, a process with a classical input x (a setting variable) that determines which of a set of sources is implemented can be represented by a set of sets of unnormalized states $\{\{p(a|x)\rho_{a|x}^A\}_a\}_x$. We refer to such a process as a multi-source. We will sometimes write $\mathbf{P} := \{\{p(a|x)\rho_{a|x}^A\}_a\}_x$. Note that when such processes arise in steering scenarios, they are also termed assemblages.

Next, we turn to processes of the measurement variety, having a single quantum input A . Three types of processes in this class are depicted in the bottom row of Fig. 2. (Note that the term measurement is used here to refer specifically to processes with no quantum output, such that there is only a retrodictive aspect to be characterized. Processes with a classical outcome that have both a quantum input and a quantum output will here be termed instruments and are discussed below.) In quantum theory, a measurement with outcome labeled by b is represented by a POVM $\{M_b^A\}_b$ where each $M_b^A \in \mathcal{M}(\mathcal{H}^A)$. The process that describes implementing this measurement and obtaining the outcome b (so that it can be represented as a process with no classical outcome) is termed an effect and is represented by the single operator $M^A \in \mathcal{M}(\mathcal{H}^A)$. We use the term multi-measurement to refer to a process with a classical input y (a setting variable) that determines which of a set of measurements is implemented. It can be represented by a set of POVMs $\mathbf{M} := \{\{M_{b|y}^A\}_b\}_y$. Note that the notion of a measurement is subsumed as a special case of a multi-measurement, where the setting variable is trivial.

Finally, we consider processes of the transformation variety. A process that has a quantum input A and a quantum output B but that has neither a classical input nor a classical output is termed a channel and is represented by a map $\mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$ that is completely positive and trace preserving [76,77]. A process of the transformation variety that also has a classical input (a setting

variable z) describes a set of channels $\{\mathcal{E}_{c|z}^{B|A}\}_z$ and will be termed a multi-channel. A process of the transformation variety that has a classical output (an outcome variable) is termed an instrument. It can be conceived of as a nondestructive measurement procedure, so that the outcome is informative not only about the quantum input (the retrodictive aspect) but also about the quantum output and the correlations between them (the transformative aspect). An instrument from $\mathcal{L}(\mathcal{H}^A)$ to $\mathcal{L}(\mathcal{H}^B)$ with classical outcome c is represented by a set of trace-nonincreasing channels denoted $\{\mathcal{F}_c^{B|A}\}_c$ that sum to a trace-preserving channel, i.e., $\sum_c \mathcal{F}_c^{B|A}$ defines a completely positive trace-preserving (CPTP) map. We use the term multi-instrument to refer to a process with a classical input z (a setting variable) that determines which of a set of instruments is implemented. We denote such a set of instruments (multi-instrument) by $\{\{\mathcal{F}_{c|z}^{B|A}\}_c\}_z$.

We note that any classical inputs (setting variables) and classical outputs (outcome variables) of a process can be reconceptualized, respectively, as quantum inputs and quantum outputs that are dephased in a particular basis. For example, a multi-measurement can be viewed as a channel that has two quantum inputs, one of which is dephased (corresponding to the setting variable), and that has a single dephased quantum output (corresponding to the outcome variable). As we will show in Sec. VIII, all of our results are invariant under this reconceptualization of the type of process.

B. Composition of quantum processes

Quantum circuits are formed via the composition of quantum processes, such as states, channels, and measurements. To discuss composition, it suffices to focus on channels, since states and effects can be treated as special cases of channels when discussing compositions, specifically:

- (i) A state ρ^A defines a channel from the trivial system to A .
- (ii) An effect M^A defines a channel $\mathcal{M}^{A|}(X^A) := \text{Tr}[M^A X^A]$ from system A to the trivial system.

We denote the operation of implementing processes on distinct systems, i.e., parallel composition of the processes, by $\otimes$. Given channels $\mathcal{E}^{B_1|A_1}: \mathcal{L}(\mathcal{H}^{A_1}) \rightarrow \mathcal{L}(\mathcal{H}^{B_1})$ and $\mathcal{D}^{B_2|A_2}: \mathcal{L}(\mathcal{H}^{A_2}) \rightarrow \mathcal{L}(\mathcal{H}^{B_2})$, their parallel composition is a bipartite channel $\mathcal{E}^{B_1|A_1} \otimes \mathcal{D}^{B_2|A_2}: \mathcal{L}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{A_2}) \rightarrow \mathcal{L}(\mathcal{H}^{B_1} \otimes \mathcal{H}^{B_2})$ (see Table II). As a special case, given states $\rho^A \in \mathcal{S}(\mathcal{H}^A)$ and $\sigma^B \in \mathcal{S}(\mathcal{H}^B)$, their parallel composition defines a bipartite state $\rho^A \otimes \sigma^B \in \mathcal{S}(\mathcal{H}^A \otimes \mathcal{H}^B)$.

We denote the operation of wiring the output of one process to the input of the other process (where the system types match), i.e., sequential composition of processes, by $\circ$. Given channels $\mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$ and $\mathcal{F}^{C|B}: \mathcal{L}(\mathcal{H}^B) \mapsto \mathcal{L}(\mathcal{H}^C)$, their sequential composition defines a

TABLE II. Different kinds of compositions (parallel and sequential) and how they are represented in quantum theory and in an ontological model thereof.

Composition	Quantum rep	Ontological rep
Parallel	$\mathcal{E}^{B_1 A_1} \otimes \mathcal{D}^{B_2 A_2}$	$p(\lambda_{B_1} \lambda_{A_1})p(\lambda_{B_2} \lambda_{A_2})$
Sequential	$\mathcal{F}^{C B} \circ \mathcal{E}^{B A}$	$\sum_{\lambda_B} p(\lambda_C \lambda_B)p(\lambda_B \lambda_A)$

new channel $\mathcal{F}^{C|B} \circ \mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^C)$ (see Table II). As a special case, the application of a channel $\mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$ to a state $\rho^A \in \mathcal{S}(\mathcal{H}^A)$ defines a new state $\mathcal{E}^{B|A} \circ \rho^A := \mathcal{E}^{B|A}(\rho^A) \in \mathcal{S}(\mathcal{H}^B)$. We will also use $\circ$ to denote more general types of “wiring together” of processes. For instance, a channel $\mathcal{E}^{B|A}$ from A to B can be slotted into a comb [78] $\mathcal{C}^{A|B}$ with output A and input B [as in Fig. 6(d) below], and we denote the resulting closed circuit by $\mathcal{C}^{A|B} \circ \mathcal{E}^{B|A}$.

Discarding (trace and partial trace) can be regarded as composition with the deterministic effect $u^{|A}: A \rightarrow I$, where I is the trivial system. Consequently, the partial trace is a special case of sequential composition. For a bipartite state $\rho^{A_1 A_2}$, one has $u^{|A} \circ \rho^{A_1 A_2} = \text{Tr}_{A_1}[\rho^{A_1 A_2}]$.

Consider a quantum circuit built up from the parallel and sequential composition of a set of quantum processes. Quantum theory provides a prescription for computing the operational statistics generated by any such a circuit. Here and throughout this article, we use the term operational statistics (or simply statistics) to refer to the conditional probability distribution over the outcome variables in the circuit given the setting variables [see Fig. 3 and Eq. (1) for an example]. Moreover, we refer to any circuit that is the sequential composition of a process of the preparation variety with a process of the measurement variety as a prepare-measure circuit or scenario. We refer to any circuit

that is the sequential composition of a process of the preparation variety with a process of the transformation variety, followed by a process of the measurement variety, as a prepare-transform-measure circuit or scenario. The same terminological convention applies to circuits on multipartite rather than unipartite systems.

C. Classical explainability of the statistics generated by a set of quantum circuits

As noted in the Introduction, we take classical explainability of operational statistics to mean the possibility of reproducing these operational statistics in a generalized-noncontextual ontological model [9,18].

Here, we wish to conceptualize quantum theory as an operational theory wherein one has quotiented the set of laboratory procedures of a given type with respect to an operational equivalence relation [79]. (This is the standard conceptualization in the field of quantum information, where a state is a density operator, a channel is a completely positive trace-preserving map, and a measurement outcome is a POVM element.) We digress briefly to explain the distinction between the unquotiented and quotiented operational theories. In an unquotiented operational theory, a given procedure (preparation, measurement, transformation, etc.) is conceptualized as a list of laboratory instructions [18], so that the theoretical description includes details that can be irrelevant to the operational statistics obtained in circuits within which the procedure is embedded. Despite their irrelevance to the operational statistics, such details, which are termed contexts, might still be relevant to the ontological representation of the procedure in a contextual ontological model [18]. Sameness of operational statistics defines an equivalence relation over procedures. A quotiented operational theory is one that includes only operational equivalence classes of procedures and does not include the details (i.e., contexts) that distinguish the procedures within a given class. For example, an equivalence class of preparation procedures in quantum theory corresponds to a density operator.

Conceptualizing quantum theory as a quotiented operational theory introduces a subtlety regarding how to understand the notion of classical explainability, namely, that the definition of a generalized-noncontextual ontological model was originally given only for unquotiented operational theories. Because a quotiented operational theory has quotiented away the notion of context that would be leveraged in a context-dependent ontological representation, it becomes a category mistake to talk about context dependence, and therefore, it becomes inappropriate to define classical explainability of such a theory as the possibility of reproducing its operational statistics in terms of a generalized-noncontextual ontological model.

What is required is a translation of the notion of classical explainability of unquotiented operational theories described above into the language of quotiented operational

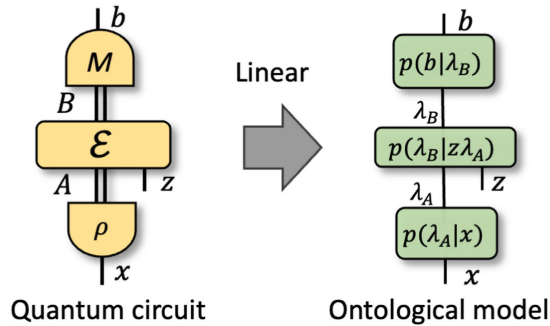


FIG. 3. A prepare-transform-measure circuit (left) and an ontological representation of it (right). The circuit is classically explainable if and only if there exists a linear and diagram-preserving ontological representation taking each quantum system to a classical random variable and each quantum process to a substochastic matrix. The map must, moreover, respect the circuit’s topology and reproduce the quantum predictions.

theories. Such a translation was presented in Refs. [9,13]. It is this: A quotiented operational theory is classically explainable if and only if its predictions can be reproduced by an ontological theory that is obtained from the quotiented operational theory by a linear and diagram-preserving ontological representation map, where each operational process is mapped to a substochastic map.

Consequently, this is our notion of classical explainability:

Definition 1. The statistics generated by a set of quantum circuits of a given structure are said to be classically explainable if and only if they admit of a linear and diagram-preserving ontological representation; that is, if and only if there is a linear and diagram-preserving map from the quantum systems appearing in the set of circuits to random variables, and from the quantum processes appearing in the set of circuits to substochastic matrices such that the map preserves the statistical predictions [9].

Note that we will sometimes simply say that the set of circuits itself (rather than its statistics) is classically explainable, and we will often consider situations where the set of circuits in question contains only one element, in which case, we will speak of the classical explainability of a circuit rather than of a set of circuits (an example of this follows). This notion of classical explainability of a set of circuits is equivalent to the notion defined by the possibility of a generalized-noncontextual ontological model of an unquotiented version of the set of circuits [9], the motivations for which were discussed in the Introduction.

This definition relies on an abstract notion of a linear and diagram-preserving (or functorial [80]) map from quantum processes in the circuit(s) to substochastic matrices over random variables. The advantage of this definition is that it is simple to state, simple to visualize (see Fig. 3), and fully general, in the sense that it applies to arbitrary circuits. The disadvantage is that writing it down more formally and explicitly requires a diagrammatic formalism like that presented in Ref. [9]. As many readers may not be familiar with this diagrammatic notation, we will not repeat the general definitions here, but will instead give only the explicit form for specific examples (for which standard algebraic notations suffice).

As a first illustrative example, imagine one implements an experiment with a prepare-transform-measure circuit of the sort shown in Fig. 3, consisting of a multi-state $\{\rho_{\cdot|x}^A\}_x$, multi-channel $\{\mathcal{E}_{\cdot|z}^{B|A}\}_z$, and measurement $\{M_b^B\}_b$.

The operational statistics one can generate within such a circuit are

$$p(b|xz) = \text{Tr} \left[M_b^B \mathcal{E}_{\cdot|z}^{B|A} (\rho_{\cdot|x}^A) \right]. \quad (1)$$

A diagram-preserving ontological representation for such a circuit provides a realist explanation of these statistics in terms of underlying spaces of ontic states for

systems A and B that encapsulate the fundamental properties of the system in question and that ultimately determine the outcomes of measurements. In particular, it associates with each system S in the circuit an ontic state space Λ_S stipulating the set of all possible ontic states λ_S . (We assume this set is finite for simplicity of the presentation.) Each state $\rho_{\cdot|x}^A$ is associated with a probability distribution over Λ_A , namely, $p(\lambda_A|x)$, where $p(\lambda_A|x) \geq 0$ and $\sum_{\lambda_A} p(\lambda_A|x) = 1$. Each channel $\mathcal{E}_{\cdot|z}^{B|A}$ is represented by a stochastic matrix, the elements of which are conditional probabilities $p(\lambda_B|z\lambda_A)$, giving the probability of transitioning from ontic state λ_A to λ_B ; this stochastic matrix must satisfy $p(\lambda_B|z\lambda_A) \geq 0$ for all $(\lambda_A, \lambda_B) \in \Lambda_A \times \Lambda_B$ and $\sum_{\lambda_B} p(\lambda_B|z\lambda_A) = 1$ for all $\lambda_A \in \Lambda_A$. Each measurement effect M_b^B is represented by a response function $p(b|\lambda_B)$ describing the probability of obtaining outcome b given the ontic state λ_B ; each response function satisfies $1 \geq p(b|\lambda_B) \geq 0$, and for a set of effects $\{M_b^B\}_b$ summing to the identity (and so forming a valid measurement), one must have $\sum_b p(b|\lambda_B) = 1$ for all $\lambda_B \in \Lambda_B$.

The requirement that the ontological model reproduces the empirical predictions of quantum theory for this prepare-transform-measure scenario of Fig. 3 implies that

$$\begin{aligned} p(b|xz) &= \text{Tr} \left[M_b^B \mathcal{E}_{\cdot|z}^{B|A} (\rho_{\cdot|x}^A) \right] \\ &= \sum_{\lambda_A \lambda_B} p(b|\lambda_B) p(\lambda_B|z\lambda_A) p(\lambda_A|x). \end{aligned} \quad (2)$$

Thus far, we have described the implications of the ontological representation being diagram preserving. We now consider the additional conditions implied by the linearity of the representation. First of all, there are generic constraints among the ontological representations of the states in the circuit, since the mapping from the set of states $\{\rho_{\cdot|x}^A\}_x$ to their ontological representations $\{p(\lambda_A|x)\}_x$ must be linear. So if an identity such as

$$\sum_x \alpha_x \rho_{\cdot|x}^A = \mathbb{0}^A \quad (3)$$

holds among the states (where here and elsewhere in the article, $\mathbb{0}$ denotes the zero operator or zero channel and $\{\alpha_x\}$ is a set of real numbers), the corresponding identity must hold among the ontological representations of these states, namely,

$$\sum_x \alpha_x p(\lambda_A|x) = 0, \quad \forall \lambda_A. \quad (4)$$

We will call constraints of the type Eq. (3) operational identities, and constraints of the type Eq. (4) ontological identities.

Similarly, any operational identity among transformations, such as

$$\sum_z \alpha_z \mathcal{E}_{\cdot|z}^{B|A} = \mathbb{0}^{B|A}, \quad (5)$$

implies the corresponding ontological identity

$$\sum_z \alpha_z p(\lambda_B | z \lambda_A) = 0, \quad \forall \lambda_A, \lambda_B, \quad (6)$$

and any operational identity

$$\sum_b \alpha_b M_b^B = \mathbb{0}^B \quad (7)$$

among effects implies the corresponding ontological identity

$$\sum_b \alpha_b p(b | \lambda_B) = 0, \quad \forall \lambda_B. \quad (8)$$

Linearity of an ontological representation, moreover, imposes constraints on how the compositions of different processes are represented.

Consider the composition of the multi-state and multi-channel from Fig. 3, yielding a new multi-state $\{\tilde{\rho}_{\cdot|zx}^B\}_{zx}$ depicted in Fig. 4(a), and consider the composition of the measurement and multi-channel from Fig. 3, yielding a new multi-measurement $\{\tilde{M}_{b|z}^A\}_z$ depicted in Fig. 4(b). The corresponding ontological representations of these

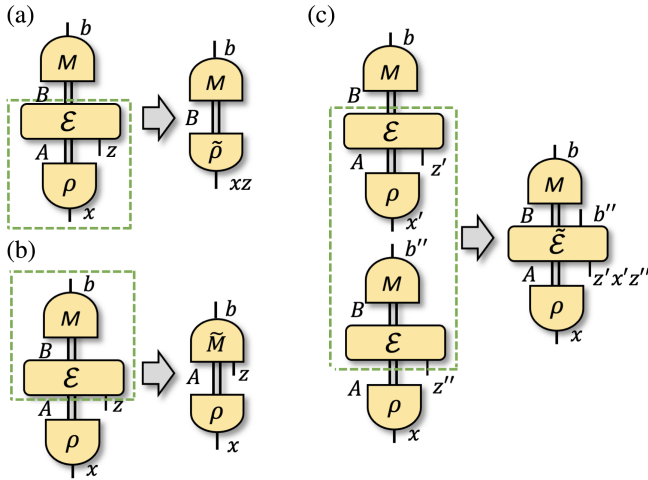


FIG. 4. In the prepare-transform-measure circuit, one can define different composite processes. (a) Boxing the multi-channel and the multi-state defines a new multi-state on $\mathcal{H}^B$, $\tilde{\rho}_{\cdot|zx}^B := \mathcal{E}_{\cdot|z}^{B|A}(\rho_{\cdot|x}^A)$. (b) Boxing the multi-channel and the measurement defines a multi-measurement on $\mathcal{H}^A$, $\tilde{M}_{b|z}^A := [\mathcal{E}_{\cdot|z}^{B|A}]^\dagger(M_b^B)$. (c) More subtly, boxing the (composite) multi-state $\{\tilde{\rho}_{\cdot|zx}^B\}_{zx}$ with the (composite) multi-measurement $\{\tilde{M}_{b|z}^A\}_z$ defines a measure-and-prepare multi-instrument from $\mathcal{H}^A$ to $\mathcal{H}^B$ with $\tilde{\mathcal{E}}_{b''|z'x'}^{B|A}(\cdot) := \tilde{\rho}_{\cdot|z'x'}^B \text{Tr}[\tilde{M}_{b''|z''}^A(\cdot)]$.

compositions are

$$\tilde{p}(\lambda_B | xz) := \sum_{\lambda_A} p(\lambda_B | z \lambda_A) p(\lambda_A | x), \quad (9a)$$

$$\tilde{p}(b | z \lambda_A) := \sum_{\lambda_B} p(b | \lambda_B) p(\lambda_B | z \lambda_A). \quad (9b)$$

Suppose that the operational identities that the new processes satisfy are

$$\sum_{zx} \alpha_{zx} \tilde{\rho}_{\cdot|zx}^B = \mathbb{0}^B, \quad (10a)$$

$$\sum_{bz} \alpha_{bz} \tilde{M}_{b|z}^A = \mathbb{0}^A; \quad (10b)$$

linearity then implies that the same identities must also hold for their ontological representations, namely,

$$\sum_{zx} \alpha_{zx} \tilde{p}(\lambda_B | xz) = 0, \quad \forall \lambda_B, \quad (11a)$$

$$\sum_{bz} \alpha_{bz} \tilde{p}(b | z \lambda_A) = 0, \quad \forall \lambda_A. \quad (11b)$$

A more subtle kind of operational identity that one must also consider [81] is the type illustrated in Fig. 4(c). Here, a multi-state $\{\tilde{\rho}_{\cdot|zx}^B\}_{zx}$ on system B [defined in Fig. 4(a)] is composed with a multi-measurement $\{\tilde{M}_{b|z}^A\}_z$ on system A [defined in Fig. 4(b)], in a manner giving rise to a multi-instrument from system A to B , namely,

$$\tilde{\mathcal{E}}_{b''|z'x'}^{B|A}(\cdot) := \tilde{\rho}_{\cdot|z'x'}^B \text{Tr}[\tilde{M}_{b''|z''}^A(\cdot)]. \quad (12)$$

(Note that this type of composition operation can be considered as an instance of either parallel or sequential composition.) The resulting multi-instrument, i.e., $\tilde{\mathcal{E}}_{b''|z'x'}^{B|A}(\cdot)$, is of the same type as the given multi-channel $\{\mathcal{E}_{\cdot|z}^{B|A}\}_z$, so there can be new operational identities of the form

$$\sum_z \alpha_z \mathcal{E}_{\cdot|z}^{B|A} + \sum_{b''x'z''} \alpha_{b''x'z''} \tilde{\mathcal{E}}_{b''|z'x'}^{B|A} = \mathbb{0}^{B|A}, \quad (13)$$

which implies the corresponding ontological identity

$$\sum_z \alpha_z p(\lambda_B | z \lambda_A) + \sum_{b''x'z''} \alpha_{b''x'z''} \tilde{p}(\lambda_B | x'z') \tilde{p}(b'' | \lambda_A z'') = 0 \quad (14)$$

for all λ_A and λ_B . Of course, this operational identity is nontrivial [i.e., not implied by the operational identities

from Eq. (5)] only if $\alpha_z \neq 0$ for some z while $\alpha_{b''x'z'z''} \neq 0$ for some $b''x'z'z''$.

If we move beyond prepare-transform-measure scenarios, the prescription is similar, but one must consider all possible ways of composing together the processes within the circuit and then demand that the operational identities on all of these effective processes are preserved in the ontological representation, in a manner analogous to what was done above. We again refer the reader to Ref. [9] for a diagrammatic framework that greatly simplifies the explicit treatment of general circuits.

It is worth emphasizing that assessments of classical explainability are always given relative to a circuit, that is, relative to a factorization of systems into subsystems and a hypothesis about the causal structure. This follows from two consequences of diagram preservation: the fact that an ontological model associates an ontic state space to each system in the given circuit and the fact that the stochastic matrices in the model must be wired together with the same connectivity as in the quantum circuit. Although this choice is quite innocuous in simple cases like a prepare-measure scenario or a Bell scenario [82], it is a bit more substantive in general. We will see this in Sec. V when we discuss the representation of bipartite systems. The reader can also find more discussion of this point in Refs. [10,70].

D. Two notions of variability in the identity of a process

There are two ways in which one can imagine variability in the identity of a process within a circuit. On the one hand, this variability might be induced by the variability in the value of a setting variable and/or in the value of an outcome variable. This is a variability that gets explored physically across different runs of the experiment. On the other hand, we sometimes wish to contemplate a theoretical notion of variability. For instance, for a process that has neither classical inputs nor classical outputs (such as a single state or transformation), we might wish to consider all possible choices for this process according to quantum theory. We refer to the former sort of variability as realized variability and the latter as theoretical variability.

In the following, when we consider a specific quantum circuit, there will be some realized variability within that circuit for a given process. This type of variability is considered part of the specification of the circuit and is captured by the ranges of the classical setting and outcome variables attached to that process.

However, in our assessments of classical explainability of a given type of process, we will often consider all the processes that are dual to it (in a sense to be specified below). The variability in the dual process is of the theoretical variety. We are considering what quantum theory predicts for every possible choice of this dual process. For the case where the process is a state, multi-state, source, or multi-source, the dual process is an effect. Quantum theory stipulates the probability assigned to all effects, and these

probability assignments do not depend on which measurement, or multi-measurement, this effect is considered part of. Similarly, for the case where the process is a measurement or a multi-measurement, the dual process is a state. Quantum theory stipulates the probability generated by all states, and these probabilities do not depend on what source or multi-source this state is considered a part of.

III. DEFINING THE CLASSICAL-NONCLASSICAL DIVIDE FOR AN INDIVIDUAL QUANTUM PROCESS

When should a given quantum process be deemed nonclassical? Intuitively, it is nonclassical if and only if (i) it can be used in some quantum circuit to realize operational statistics that are not classically explainable (in the sense of Definition 1), and (ii) it is implicated in the lack of classical explainability. We begin by clarifying what we mean by “being implicated in the lack of classical explainability.” Essentially, if a circuit realizes statistics that are not classically explainable independent of the identity of the process in question, the process is not implicated in the lack of classical explainability.

Consider the question of whether a given multi-source is nonclassical. Suppose one has the circuit in Fig. 5(a), where the multi-source on system A appears in yellow, and suppose that an ancilla system labeled B is also prepared in some state (in blue), and the composite system and ancilla are subjected to a joint measurement. In this case, it is possible to observe a failure of classical explainability

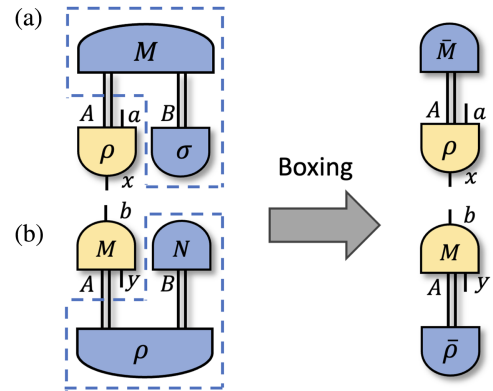


FIG. 5. Left: probing schemes (in blue) that can generate statistics that are not classically explainable independent of the identity of the given process of interest (in yellow). In (a), the process of interest is a multi-source. In (b), the process of interest is a multi-measurement. Right: treating these same circuits in a manner that coarse grains over all processes involving systems that are not connected directly to the process of interest leads to probing schemes (in blue) such that whether the statistics can be explained classically now always depends nontrivially on the identity of the process of interest. In (a), the effective process is simply an effect $\bar{M}_A$; in (b), the effective process is simply a state $\bar{\rho}_A$.

regardless of the identity of the multi-source on the system. This can be done, for instance, simply by taking the multipartite measurements to be factorized across the system-ancilla divide and by choosing any set of states and any set of effects on the ancilla sufficient to generate statistics that are not classically explainable (in the sense of Definition 1).

Similarly, consider whether a given multi-measurement is nonclassical. Suppose one has the circuit in Fig. 5(b), where the multi-measurement appears in yellow, and suppose that the composite of the system and ancilla is prepared in some joint state (in blue), and subsequently a factorizing measurement is made, with some local measurement implemented on the ancilla. Again, it is possible to observe the failure of classical explainability regardless of the identity of the multi-measurement on the system: Simply take the joint state to be factorizing across the system-ancilla divide, and choose any set of states and set of effects on the ancilla is sufficient to generate statistics that are not classical explainable (in the sense of Definition 1).

In both cases, then, one cannot conclude that the process in question is implicated in the fact that the overall circuit admits no classical explanation.

To ensure that the lack of classical explainability of the circuit implies the nonclassicality of the process of interest, it suffices to stipulate that for any circuit that includes the process of interest, one must box together (i.e., compose together) all processes that act on ancillary quantum systems. In other words, this circuit should not introduce additional causal structure that is internal to the probing part of the circuit.

In the above examples, this entails boxing together the blue processes to define an effective process: In Fig. 5(a), one must compose the effect M^{AB} with the state σ^B to define an effective effect $\bar{M}^A := M^{AB} \circ \sigma^B$ [i.e., $\text{Tr}[\bar{M}^A(\cdot)] := \text{Tr}[M^{AB}(\cdot \otimes \sigma^B)]$] on system A alone. In Fig. 5(b), one must compose the effect N^B with the state ρ^{AB} to define an effective state $\bar{\rho}^A := N^B \circ \rho^{AB}$ [i.e., $\bar{\rho}^A := \text{Tr}[(\mathbb{1}^A \otimes N^B)\rho^{AB}]$] on system A alone. By boxing processes in this manner, one ensures that any nonclassicality in the circuit comes from constraints pertaining to the system of interest (here, system A), as opposed to any ancilla systems internal to the probing part of the circuit (here, system B).

Equivalently, one can simply define nonclassicality of a given process in terms of whether or not it can generate statistics that are not classically explainable when embedded in a circuit that does not introduce any new quantum systems. That is, the circuit cannot contain any quantum systems not connected directly to the process itself (for instance, it cannot contain a system such as B in Fig. 5). This is consistent with the fact that noncontextuality is always defined relative to some specified systems. Consequently, one wishes to consider the possibility of classical explanations for the systems in question (those connected to the given process) and no others, and the stipulation just given is a simple heuristic to guarantee this.

We formalize all of this by defining the notion of a dual process to P .

Definition 2. For a given type of quantum process P , a dual process to P is a quantum process that has a quantum input corresponding to every quantum output of P and a quantum output corresponding to every quantum input of P , and that has no other quantum inputs or outputs.

If one connects (that is, one “wires together”) each quantum output of P to the corresponding quantum input of the dual process and each quantum input of P to the corresponding quantum output of the dual process, one obtains a circuit with no open quantum wires and with which one can associate a probability (for each value of the setting variables and outcome variables of the process P). It follows that a dual process defines a linear functional on the process P . We will refer to this wiring together of D and P to form $D \circ P$ as composing D with P .

We depict examples of dual processes in Fig. 6. For the process of the preparation variety on A , i.e., a state, multi-state, source, or multi-source on system A , the dual process is an effect on A . Similarly, for the process of the measurement variety on A , i.e., an effect, measurement, or multi-measurement on system A , the dual process is a state on A .

Throughout the paper, we treat the variability in the states and effects of a dual process as theoretical variability (as discussed in Sec. II D). Consequently, dual processes in our treatment do not have classical inputs or classical outputs. As discussed in Sec. II D, there are many different ways of ranging over all dual processes of a given type through physical variability (i.e., by the variability in the values of classical inputs and classical outputs), but the definition of classicality is independent of the choice of how to do so. This is discussed further in Appendix C.

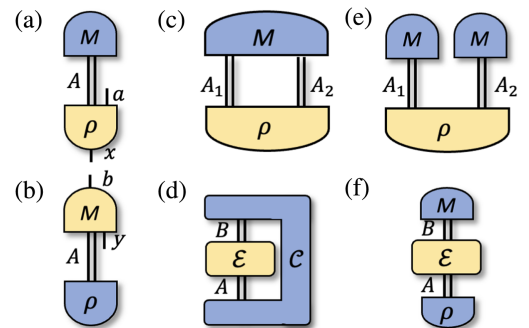


FIG. 6. In (a)–(d), we depict examples of various types of quantum processes (in yellow) and their associated dual processes (in blue). In (a), the process is a multi-source on a single system. In (b), it is a multi-measurement on a single system. In (c), it is a state on a bipartite system. In (d), it is a channel. In (e) and (f), the processes are the same as in (c) and (d), respectively, but we depict the associated dual processes when the latter are factorizing.

For a state on a multipartite system, the dual process is an effect on that multipartite system, as depicted in Fig. 6(c). The same is true if one considers any process of the preparation variety, i.e., a multipartite multi-state, source, or multi-source. For processes with both nontrivial quantum inputs and nontrivial quantum outputs, the dual process is a comb [78,83], as depicted in Fig. 6(d).

With these notions in hand, we can present our proposed notion of classicality for an individual process.

Definition 3. A quantum process is classical if and only if the statistics generated by the set of circuits that contract the process of interest with any dual process (i.e., where the set ranges over all choices of dual process) are classically explainable (in the sense of Definition 1).

It is useful to provide a slightly more formal version of this definition. Suppose $\mathcal{T}$ denotes a variety of a process, such as a preparation on A , a preparation on AB , a measurement on A , a transformation from A to B , etc. For a process P of type $\mathcal{T}$, let the set of all processes of the dual type be denoted $\text{AllDuals}(\mathcal{T})$, and recall that the circuit one obtains by contracting a process P with a dual process D is denoted $D \circ P$. Relative to these notational conventions, the definition can be summarized as follows: A quantum process P is classical if and only if the statistics generated by the set of circuits $\{D \circ P | D \in \text{AllDuals}(\mathcal{T})\}$ are classically explainable.

Note that this basic definition generalizes naturally to theories beyond quantum theory, that is, to arbitrary GPTs. However, in this paper, we focus specifically on quantum theory. As a consequence, we can simplify the definition considerably.

It turns out that it is sufficient to focus on a special class of dual processes. To define this class, we introduce the notion of a factorizing process.

Definition 4. A process is termed factorizing if it consists of a product of local effects on its quantum inputs and a product of local states on its quantum outputs.

A few examples serve to illustrate the idea. A state on a bipartite system AB is factorizing if it is a product of a state on A and a state on B . A process with quantum inputs A and B and a single quantum output C is factorizing if it is a product of an effect on A and an effect on B , composed sequentially with the preparation of a state on C . A comb with quantum output A and quantum input B is factorizing if it consists of a state on A and an effect on B .

With this definition, we prove (in Appendix D) the following:

Theorem 1. A quantum process is such that there exists some set of dual processes that it can be contracted with to obtain statistics that are not classically explainable, if and only if there exists some set of factorizing dual processes for which this is the case.

Consequently, we can restrict our attention to factorizing dual processes. For instance, for a bipartite state, it suffices to consider product effects, as depicted in Fig. 6(e). For a

channel, it suffices to consider factorizing combs, as depicted in Fig. 6(f). With this result in hand, we can provide an alternative definition to Definition 3 of the classical-nonclassical divide for a single quantum process as follows.

Definition 3'. A quantum process is classical if and only if the statistics generated by the set of circuits that contract the process of interest with any factorizing dual process (i.e., where the set ranges over all choices of factorizing dual process) are classically explainable (in the sense of Definition 1).

This is the central definition of our manuscript.

IV. PROCESSES OF THE PREPARATION OR MEASUREMENT VARIETY

Let us start by reviewing the form of a classical explanation (an ontological model) for the statistics in the simplest prepare-measure scenario. We will consider the most general circuit in such a scenario, with a multi-source on the preparation side and a multi-measurement on the measurement side.

In this section, we will drop the superscript specifying the quantum system since only one system is involved.

A. Classical explainability of a prepare-measure scenario

In the prepare-measure scenario, we consider a multi-source $\mathbf{P} := \{\{p(a|x)\rho_{a|x}\}_a\}_x$ and a multi-measurement $\mathbf{M} := \{\{M_{b|y}\}_b\}_y$, which, when contracted together (as depicted in Fig. 7), generate the statistics

$$p(ab|xy) = p(a|x)\text{Tr}[M_{b|y}\rho_{a|x}]. \quad (15)$$

In this simple context, the only constraints implied by diagram preservation in this circuit are that states $\rho_{a|x}$ get represented as probability distributions $p(\lambda|ax)$ [the subnormalized state $p(a|x)\rho_{a|x}$ gets represented as $p(a\lambda|x) := p(a|x)p(\lambda|ax)$], that effects get represented as response functions $p(b|y\lambda)$, and that one can reproduce the quantum statistics as

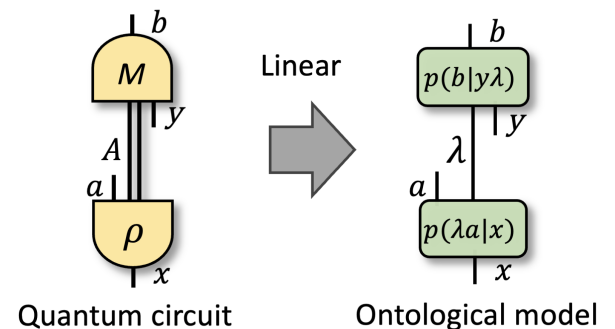


FIG. 7. A prepare-measure circuit (left) and an ontological model for it (right).

$$p(ab|xy) = \sum_{\lambda} p(b|y\lambda)p(a\lambda|x). \quad (16)$$

The states and effects in such a circuit satisfy operational identities. We will identify these operational identities by their coefficients, and so we can denote the set of all operational identities for the states and that for the effects, respectively, by

$$\mathcal{O}(\mathbf{P}) := \{ \{ \alpha_{ax} \}_{ax} \mid \sum_{ax} \alpha_{ax} p(a|x) \rho_{a|x} = 0 \}, \quad (17a)$$

$$\mathcal{O}(\mathbf{M}) := \{ \{ \alpha_{by} \}_{by} \mid \sum_{by} \alpha_{by} M_{b|y} = 0 \}. \quad (17b)$$

If one further demands that the ontological representation is linear, the conditional probabilities $p(b|y\lambda)$ and $p(a\lambda|x)$ must satisfy the ontological identities

$$\sum_{ax} \alpha_{ax} p(a\lambda|x) = 0 \quad \forall \{ \alpha_{ax} \}_{ax} \in \mathcal{O}(\mathbf{P}), \quad (18a)$$

$$\sum_{by} \alpha_{by} p(b|y\lambda) = 0 \quad \forall \{ \alpha_{by} \}_{by} \in \mathcal{O}(\mathbf{M}) \quad (18b)$$

for all λ . Consequently, this circuit is classically explainable (i.e., admits a linear and diagram-preserving ontological representation) if and only if the statistics can be reproduced by Eq. (16) under the constraint of Eq. (18).

B. The classical-nonclassical divide for multi-sources

Next, we characterize what states, multi-states, sources, and multi-sources are classical. Because all of these are special cases of the single concept of a multi-source, it suffices to define classicality for a multi-source. This is just a special case of Definition 3':

Definition 5. A multi-source is classical if and only if the statistics generated by the set of circuits that contract it with any effect (i.e., where the set ranges over all possible choices of effects) are classically explainable (in the sense of Definition 1).

The contraction of a multi-source with an effect results in a circuit of the prepare-measure form, as depicted in Fig. 6(a).

Precisely which multi-sources are classical is characterized by the following theorem, which refers to the notions of frame and dual frame. See Appendix A for a review of these.

Theorem 2. For a multi-source

$$\mathbf{P} := \{ \{ p(a|x) \rho_{a|x} \}_a \}_x,$$

the following statements are equivalent:

- (0) It is classical (in the sense of Definition 5).

- (1) Each of the subnormalized states $p(a|x) \rho_{a|x}$ can be decomposed as

$$p(a|x) \rho_{a|x} = \sum_{\lambda} p(a\lambda|x) \sigma_{\lambda}, \quad \forall a, x \quad (19)$$

for a set of normalized states $\{ \sigma_{\lambda} \}_{\lambda}$ and some conditional probability distribution $p(a\lambda|x)$ satisfying

$$\sum_{ax} \alpha_{ax} p(a\lambda|x) = 0, \quad \forall \{ \alpha_{ax} \}_{ax} \in \mathcal{O}(\mathbf{P}) \quad (20)$$

for all λ , where $\mathcal{O}(\mathbf{P})$ is defined in Eq. (17a).

- (2) Each of the normalized states $\rho_{a|x}$ can be decomposed within a frame representation on $V = \text{Span}(\{ \rho_{a|x} \}_{ax})$ as

$$\rho_{a|x} = \sum_{\lambda} \text{Tr}[H_{\lambda} \rho_{a|x}] \sigma_{\lambda}, \quad \forall a, x, \quad (21)$$

where the frame is a fixed set of density operators $\{ \sigma_{\lambda} \}_{\lambda}$, and the dual frame is a set of (possibly nonpositive) Hermitian operators, i.e., $\{ H_{\lambda} \}_{\lambda} \subset \text{Herm}(\mathcal{H})$ satisfying $\sum_{\lambda} H_{\lambda} = \mathcal{P}_V(\mathbb{1})$ and $\text{Tr}[H_{\lambda} \rho_{a|x}] \in [0, 1]$ for all a, x, λ , with $\mathcal{P}_V$ being the superoperator that projects onto the operator space V . This condition is depicted in Fig. 8.

Remark. We introduce the superoperator $\mathcal{P}_V$ to reflect the fact that $\sum_{\lambda} H_{\lambda}$ need not equal the identity operator $\mathbb{1}$ on $\mathcal{H}$. [A sufficient condition for $\sum_{\lambda} H_{\lambda} = \mathbb{1}$ is that V coincides with the operator space $\text{Herm}(\mathcal{H})$.]

As an example, the multi-state $\rho_{\cdot|x} = \{ |0\rangle\langle 0|, |+\rangle\langle +| \}$ is classical with a frame representation of the form of Eq. (21) with $\sigma_0 = |0\rangle\langle 0|$, $\sigma_1 = |+\rangle\langle +|$ and $H_0 = \frac{4}{3}|0\rangle\langle 0| - \frac{2}{3}|+\rangle\langle +|$, $H_1 = \frac{4}{3}|+\rangle\langle +| - \frac{2}{3}|0\rangle\langle 0|$. However, in this case, $\sum_{\lambda} H_{\lambda} = \mathcal{P}_V(\mathbb{1}) = \frac{2}{3}|0\rangle\langle 0| + \frac{2}{3}|+\rangle\langle +| \neq \mathbb{1}$.

The proof is given in Appendix E. We will sometimes refer to this result as a structure theorem for classical multi-sources. In a companion paper [72], we leverage this theorem to provide detailed results on the certification and quantification of the nonclassicality of general multi-sources.

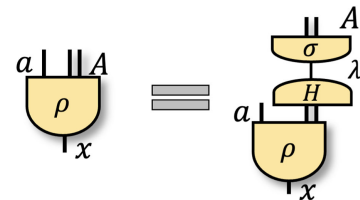


FIG. 8. A diagrammatic criterion for classicality of a multi-source $\{ \{ p(a|x) \rho_{a|x} \}_a \}_x$: a multi-source is classical if and only if it admits the frame representation as shown on the right, where the conditions on the frame and dual-frame operators are given in Theorem 2.

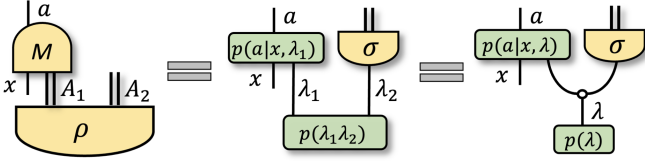


FIG. 9. A steering experiment defines a multi-source $\{p(a|x)\rho_{a|x}\}_x$ often termed an assemblage. An assemblage is said to be unsteerable (or free) if and only if it admits a decomposition in terms of local operations and shared randomness, as $\sum_{\lambda_1, \lambda_2} p(a|x\lambda_1)p(\lambda_1\lambda_2)\sigma_{\lambda_2}$, as shown in the middle figure, or equivalently, as $\sum_{\lambda} p(\lambda)p(a|x\lambda)\sigma_{\lambda}$, as shown in the right-hand-side figure, where the dot represents the copying operation.

We can also derive some immediate insights from this theorem. First, consider multi-sources that arise in quantum steering experiments [84,85]. As shown on the left-hand side of Fig. 9, a pair of systems, A_1 and A_2 , is prepared in an entangled state, and a multi-measurement is implemented on A_1 . This yields an effective process that is a multi-source on A_2 . Such a multi-source always satisfies $\sum_a p(a|x)\rho_{a|x} = \sigma$ for some state σ independent of x , and is typically termed an assemblage [86]. An assemblage is said to be unsteerable [84] if its unnormalized states can be decomposed as

$$p(a|x)\rho_{a|x} = \sum_{\lambda} p(\lambda)p(a|x\lambda)\sigma_{\lambda} \quad (22)$$

for some set $\{\sigma_{\lambda}\}_{\lambda}$ of density operators on A_2 , some probability distribution $p(\lambda)$, and some conditional probability distribution $p(a|x\lambda)$, as depicted in Fig. 9. This is often referred to as a local hidden state model of the assemblage. (Note that in the resource theory of local operations and shared randomness [2–4,87], the unsteerable assemblages are the free resources when one considers processes corresponding to a steering experiment [88,89].) If an assemblage fails to satisfy the condition for being unsteerable, it is said to be steerable.

Given the definition of steerability, Theorem 2 immediately implies the following result.

Corollary 1. Every steerable assemblage is nonclassical.

Proof. We prove the contrapositive that if the multi-source defined by an assemblage is classical, then it is unsteerable. The assumption of classicality is implied by the satisfaction of condition (1) of Theorem 2. Summing Eq. (19) over the outcome index a , and using the fact that an assemblage satisfies the operational identity

$$\sum_a p(a|x)\rho_{a|x} = \sum_a p(a|x')\rho_{a|x'} \quad \text{for all } x, x'. \quad (23)$$

By Eq. (20), it follows that

$$\sum_a p(a\lambda|x) = \sum_a p(a\lambda|x'), \quad \forall x, x'. \quad (24)$$

Defining $p(\lambda|x) := \sum_a p(a\lambda|x)$, we conclude that $p(\lambda|x)$ is independent of x ; i.e., we can denote

$$p(\lambda|x) = p(\lambda), \quad \forall x. \quad (25)$$

By Bayesian inversion,

$$p(a\lambda|x) = p(a|x\lambda)p(\lambda|x), \quad (26)$$

and so Eq. (25) implies

$$p(a\lambda|x) = p(a|x\lambda)p(\lambda). \quad (27)$$

Plugging this into Eq. (19), we obtain Eq. (22), the condition for being an unsteerable assemblage. ■

In other words, an assemblage being classical (in the sense of Definition 3) implies that it admits a local hidden state model and that it is a free resource relative to local operations and shared randomness. It is worth noting that, in our framework, every steerable assemblage remains nonclassical even when it is viewed as a general multi-source, i.e., without necessarily assuming the lack of causal influence from the setting to the quantum output that is implied by the steering scenario. However, the converse is not true, as is established below in Example 2.

Theorem 2 also has consequences for the particular case of a multi-state where the states are linearly independent. For any multi-state described by a set $\{\rho_{\cdot|x}\}_x$ of linearly independent states on a single system, the set $\mathcal{O}(\mathbf{P})$ of operational identities it satisfies is empty, so [by Eq. (20)] there are no constraints on the set of distributions $\{p(\lambda|x)\}_x$ that are the ontological representations of these states. Consequently, one can always find a decomposition as in Eq. (19) (where a is trivial) by defining $\sigma_{\lambda} := \rho_{\cdot|\lambda}$ and $p(\lambda|x) = \delta_{x,\lambda}$. So one has the following result (also noted in Ref. [70]).

Corollary 2. A multi-state consisting of a set of linearly independent states on a single system is classical.

From this, one can prove that any multi-source defined by a set of three or fewer states on a single system is also classical. (The classical explainability of a prepare-measure scenario involving fewer than four states was first noted in Ref. [90].) The case where the states are linearly independent is covered by Lemma 2, so it suffices to consider the case where they are linearly dependent. The single-state case is trivial. For a pair of states, linear dependence implies that the states are equal, so we again have a trivial case. For a triple of states, any possible linear dependence has the form $\rho_3 = p\rho_1 + (1-p)\rho_2$ for some $p \in [0, 1]$ (for some suitable relabeling). There is consequently a single operational identity in the set $\mathcal{O}_p$. But in this case, one can find a decomposition as in Eq. (19), where a is trivial, where $\lambda := \{1, 2\}$, where $\sigma_1 := \rho_1$ and $\sigma_2 := \rho_2$, and where $p(\lambda|x) = \delta_{x,\lambda}$ for $x \in \{1, 2\}$ and $p(\lambda|x) = p\delta_{\lambda,1} + (1-p)\delta_{\lambda,2}$ for $x = 3$. The last equation ensures that Eq. (20) is satisfied and hence that condition (1) of Theorem 2 holds.

This argument generalizes to give the following sufficient condition for classicality of a set of states.

Corollary 3. If a multi-state consists of a set of states that fits inside a simplex within the quantum state space, then it is classical.

This is proven in Appendix G.

For a set of distinct pure states, linear dependence is not only necessary for nonclassicality (per Corollary 2), but also sufficient.

Corollary 4. A multi-state consisting of a set of distinct pure states on a single system is classical if and only if the states are linearly independent.

The proof is in Appendix H.

So, for example, any set of five pure qubit states is nonclassical, as are any four pure states lying within a plane of the Bloch ball. An example of such a set of four states arises in the BB84 protocol [91] and in parity-oblivious multiplexing [44].

Example 1. The multi-state $\{|0\rangle\langle 0|, |1\rangle\langle 1|, |+\rangle\langle +|, |-\rangle\langle -|\}$ is nonclassical.

Follows immediately from Theorem 2 (condition 2), the qualitative nonclassicality of a multi-state is not affected by rescaling the states contained therein by any numbers in $(0, 1]$ (this is also proven in Ref. [32]). Consequently, we infer the following result.

Example 2. The source $\{p(a)\rho^A\}_a = \{\frac{1}{4}|0\rangle\langle 0|, \frac{1}{4}|1\rangle\langle 1|, \frac{1}{4}|+\rangle\langle +|, \frac{1}{4}|-\rangle\langle -|\}$ is nonclassical.

Considered as an assemblage, one sees that it is unsteerable by noting that it has no setting variable (i.e., it is a single ensemble rather than a set of ensembles), and so Eq. (22) is satisfied trivially; this, therefore, is an example of an unsteerable assemblage that is nonclassical.

Note that the converses of Corollaries 2 and 3 do not hold; there exist multi-states that exhibit nontrivial linear dependence relations, and moreover that do not fit inside a simplex within the quantum state space, but are nevertheless classical. This is illustrated by the following example.

Example 3. Consider the multi-states

$$\left\{ \frac{1}{2} [\mathbb{1} + \eta \hat{n}_x] \right\}_x, \quad (28)$$

where $\{\hat{n}_x\}$ are unit vectors corresponding to the eight vertices of a cube inscribed in the Bloch sphere. This set is classical if and only if $\eta \leq (1/\sqrt{3})$. (See companion paper [72] for the proof.) However, these states are linearly dependent for all $\eta > 0$; moreover, it can be verified geometrically that a shrunken cube with vertices $\{(1/\sqrt{3})\hat{n}_x\}$ cannot be contained inside any tetrahedron that fits inside the Bloch sphere, and so the set of states when $\eta = (1/\sqrt{3})$ does not fit inside any simplex within the qubit state space.

This example also highlights that our notion of classical multi-states is characterized by a frame representation (as defined in Theorem 2), where the frame may be over-complete, and the associated decomposition need not be simplicial (i.e., it need not correspond to a basis).

C. The classical-nonclassical divide for multi-measurements

Next, we turn to measurements. Recall that we use the term “measurement” to refer to a device with a quantum input and an outcome variable (no quantum output), and the term “multi-measurement” to refer to such a device when it has a setting variable as well.

The definition of nonclassicality of a multi-measurement is again a special case of Definition 3’:

Definition 6. A multi-measurement is classical if and only if the statistics generated by the set of circuits where it is contracted with any state (i.e., where the set ranges over all states) are classically explainable (in the sense of Definition 1).

Precisely which multi-measurements are classical is characterized by the following theorem.

Theorem 3. For a multi-measurement $M = \{\{M_{b|y}\}_b\}_y$, the following statements are equivalent:

- (0) It is classical (in the sense of Definition 6).
- (1) Its effects can be decomposed as

$$M_{b|y} = \sum_{\lambda} p(b|y\lambda) G_{\lambda}, \quad \forall b, y \quad (29)$$

for some POVM $\{G_{\lambda}\}_{\lambda}$ and some conditional probability distribution satisfying

$$\sum_{by} \beta_{by} p(b|y\lambda) = 0, \quad \forall \{\beta_{by}\}_{by} \in \mathcal{O}(M) \quad (30)$$

for all λ , where $\mathcal{O}(M)$ is defined in Eq. (17b).

- (2) Its effects can be decomposed within a frame representation on $V = \text{Span}(\{M_{b|y}\})$ as

$$M_{b|y} = \sum_{\lambda} \text{Tr}[M_{b|y} F_{\lambda}] G_{\lambda}, \quad \forall b, y, \quad (31)$$

where $\{G_{\lambda}\}_{\lambda}$ is a POVM with $\sum_{\lambda} G_{\lambda} = \mathcal{P}_V(\mathbb{1}) = \mathbb{1}$, and $\{F_{\lambda}\}_{\lambda} \subset \text{Herm}(\mathcal{H})$ is a set of Hermitian operators satisfying $\text{Tr}[F_{\lambda}] = 1$ and $\text{Tr}[M_{b|y} F_{\lambda}] \in [0, 1]$ for all b, y, λ . This condition is depicted in Fig. 10.

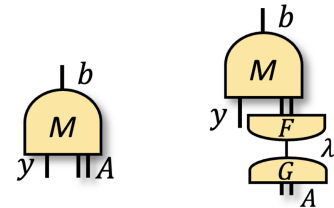


FIG. 10. A diagrammatic criterion for classicality of a multi-measurement $\{\{M_{b|y}\}_b\}_y$: a multi-measurement is classical if and only if it admits a frame representation as shown on the right, where the conditions on the frame and dual-frame operators are given in Theorem 3.

The proof of the theorem is provided in Appendix F. We will sometimes refer to this as a structure theorem for classical multi-measurements.

Remark. Here, $\sum_{\lambda} G_{\lambda} = \mathcal{P}_V(\mathbb{1}) = \mathbb{1}$ since the operator space $V = \text{Span}(\{M_{b|y}\})$ contains the identity operator (it suffices to recall that $\sum_b M_{b|y} = \mathbb{1}$). Note the difference to the situation with the dual frame $\{H_{\lambda}\}_{\lambda}$ in Theorem 2.

It is worth noting that an alternative proof of Theorem 3 is possible by leveraging Theorem 2 and Lemma 1 from Sec. VII. Essentially, a multi-measurement defines a special type of multi-source (known as a steering assemblage) via the Choi isomorphism, and Lemma 1 asserts that the verdict of nonclassicality is preserved under the Choi isomorphism. Conditions (1) and (2) in Theorem 3 are the translations through the Choi isomorphism of conditions (1) and (2) of Theorem 2.

Several interesting insights follow from this theorem, most of which are direct analogs of those in the previous section.

First, recall that a set of measurements is compatible [92] if and only if there is a single measurement $\{G_{\lambda}\}_{\lambda}$ that can be postprocessed to reproduce every measurement in the set. Representing this set of measurements as a multi-measurement $\{\{M_{b|y}\}_y\}_y$, the condition of compatibility is

$$M_{b|y} = \sum_{\lambda} p(b|y, \lambda) G_{\lambda}, \quad (32)$$

as depicted in Fig. 11. Any set of measurements for which the condition cannot be satisfied is said to be incompatible.

Recalling this, Theorem 3 immediately implies the following sufficient condition for nonclassicality.

Corollary 5. Every multi-measurement consisting of a set of incompatible measurements is nonclassical.

We prove this by demonstrating the contrapositive, that classicality implies compatibility. Specifically, classicality implies the satisfaction of condition 1 [Eq. (29)] of Theorem 3, which implies that the POVM $\{G_{\lambda}\}_{\lambda}$ defines a single measurement that can simulate all of the measurements in the given set via postprocessing [92]; hence, the set of measurements is compatible. The converse of

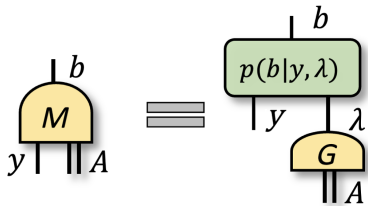


FIG. 11. A set of measurements $\{\{M_{b|y}\}_b\}_y$ is compatible if and only if there is a single measurement which can be classically postprocessed to simulate each element of the set. This implies that a multi-measurement ranging over those measurements decomposes as shown (with the single measurement denoted by $\{G_{\lambda}\}$).

Corollary 5 does not hold due to the extra constraints in Eq. (30) of condition 1; there are sets of measurements that are all compatible and yet for which the multi-measurement they define is nonetheless nonclassical.

A more exhaustive comparison between nonclassicality and incompatibility of a set of measurements is carried out in our companion paper [72].

The analog for effects of our Corollary 4 holds by an analogous proof:

Corollary 6. A measurement defined by a set of rank-1 effects is classical if and only if the effects are linearly independent.

One significant implication of this corollary is that a measurement represented by a rank-1 symmetric informationally complete POVM is classical [93].

Unlike a single state, a single measurement $\{M_b\}_b$ can be nonclassical, provided that there are nontrivial operational identities among its effects [94].

Example 4. The noisy Pentagon measurement $\{M_b^{\eta}\}_{b=0}^4$, where

$$M_b^{\eta} = \frac{1}{5} [\mathbb{1} + \eta(\cos \theta_b \sigma_x + \sin \theta_b \sigma_z)], \quad \theta_b = \frac{2\pi b}{5}, \quad (33)$$

and $\eta \in [0, 1]$ is a noise parameter (this multi-measurement was defined for $\eta = 1$ in Ref. [94]). This single measurement is nonclassical if and only if $\eta \gtrsim 0.618$ as proven in Ref. [72].

V. PROCESSES OF THE BIPARTITE PREPARATION VARIETY

Recall that a single state on a unipartite system is classical, as a consequence of Corollary 2. However, once one considers multipartite systems, even a single state can be nonclassical. This is because the subsystem structure provides new opportunities for witnessing nonclassicality (as first recognized in Ref. [70]).

When defining a linear and diagram-preserving ontological representation of a circuit that involves parallel composition of two or more systems (represented diagrammatically by a pair of wires), the assumption of diagram preservation implies that each system is associated with its own independent ontic state space. The ontic state space of the joint system is the Cartesian product of the ontic state spaces of the subsystems, an assumption that is sometimes called ontic separability.

Note that, if one instead wishes to entertain an ontological representation wherein the ontic state space of the joint system is not a Cartesian product, then one should rather represent the bipartite system in the quantum circuit as a single monolithic system—without a factorization into subsystems. In this case, the assumption of diagram preservation would not force one to assume ontic separability. Under such a choice, the joint system would be

treated as a unipartite system and represented diagrammatically by a single wire rather than a pair of wires.

Recall from Sec. II C that assessments of nonclassicality are always made relative to a given circuit, which effectively embodies a hypothesis about the causal structure. It should be clear from the previous paragraph that for a quantum system represented by a four-dimensional Hilbert space, viewing it as (i) a single monolithic system or (ii) as a pair of qubits (according to some factorization of the system into a pair of two-dimensional subsystems) constitutes two different hypotheses about the causal structure that are represented by different circuits. Indeed, the latter view, wherein a privileged factorization is specified for the full system, gives a more fine-grained perspective on the system. This is reflected in the fact that an ontological model associates with a monolithic system an ontic state space with no particular structure, while it associates with a bipartite system an ontic state space that has extra structure; it must be the Cartesian product of two ontic state spaces, one for each subsystem.

Assessments of nonclassicality may depend on which view of the system one takes. In particular, the unipartite view may lead to an assessment of classicality, while the bipartite view may lead to an assessment of nonclassicality. (Note that if the bipartite view leads to an assessment of classicality, then the unipartite view will as well.) Indeed, we will see this type dependence in the following results, where we will show that any entangled state on a bipartite system is nonclassical relative to the specified bipartition, whereas we saw in Sec. IV A (from Corollary 2, for example) that any individual state is classical when it is conceptualized as a state on a unipartite system with no specified subsystem structure.

To discuss the nonclassicality of a bipartite state, one must consider the circuit obtained by contracting it with its dual process, which is a bipartite effect, as depicted in Fig. 6(c). Given Theorem 1, however, it suffices to consider the effects that factorize across the bipartition—hence, only local measurements. Thus, it suffices to consider only the type of circuits depicted in Fig. 6(e).

A. Classical explainability of a prepare-measure scenario on a bipartite system with local measurements

As preparation for our discussion of nonclassicality of bipartite states, we consider the question of when a prepare-measure scenario involving a bipartite state and local measurements is classically explainable in the sense of Definition 1.

Specifically, we consider a circuit wherein the composite system A_1A_2 is prepared in a particular bipartite state $\mathbf{P}^{A_1A_2} := \rho^{A_1A_2}$, and then A_1 is subjected to a multi-measurement $\mathbf{M}_1^{A_1}$ while A_2 is subjected to a multi-measurement $\mathbf{M}_2^{A_2}$, where $\mathbf{M}_i^{A_i} := \{\{M_{b_i|y_i}^{A_i}\}_{b_i}\}_{y_i}$. In such an experiment depicted in Fig. 12(a), quantum theory predicts that

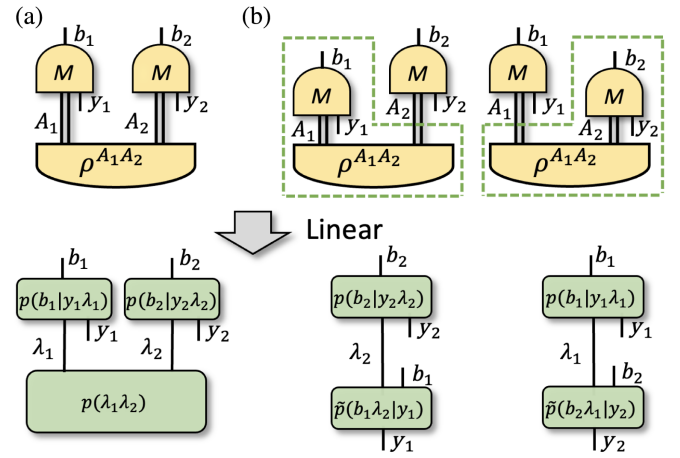


FIG. 12. (a) Local measurements on a bipartite state. (b) By boxing processes together, we can view this circuit as a set of preparations and measurements on A_2 , or as a set of preparations and measurements on A_1 . In the top of the figure, we depict the quantum circuits; in the bottom, we depict an ontological model of each circuit.

$$p(b_1b_2|y_1y_2) = \text{Tr}[\rho^{A_1A_2} M_{b_1|y_1}^{A_1} \otimes M_{b_2|y_2}^{A_2}]. \quad (34)$$

There may be operational identities among the effects on each system, namely,

$$\mathcal{O}(\mathbf{M}_1) = \{\{\alpha_{b_1y_1}\} | \sum_{b_1y_1} \alpha_{b_1y_1} M_{b_1|y_1}^{A_1} = 0\}, \quad (35a)$$

$$\mathcal{O}(\mathbf{M}_2) = \{\{\alpha_{b_2y_2}\} | \sum_{b_2y_2} \alpha_{b_2y_2} M_{b_2|y_2}^{A_2} = 0\}. \quad (35b)$$

Note that for the sake of reducing notational clutter, we will sometimes drop the system-identifying superscripts on the names of processes when these appear as arguments inside $\mathcal{O}(\cdot)$. Thus, we write $\mathcal{O}(\mathbf{M}_i)$ instead of $\mathcal{O}(\mathbf{M}_i^{A_i})$.

There may also be further operational identities in this circuit, namely, identities satisfied by the steering assemblages on A_1 and A_2 depicted in Fig. 12(b). For instance, the steering assemblage, i.e., the multi-source, on A_2 is obtained by composing the multi-measurement on A_1 with the bipartite state on A_1A_2 , and similarly for the steering assemblage on A_1 . Formally, the multi-sources on A_1 and A_2 are defined by

$$\begin{aligned} \mathbf{M}_1^{A_1} \circ \mathbf{P}^{A_1A_2} &:= \{\{\tilde{\rho}_{b_1|y_1}^{A_2}\}_{b_1}\}_{y_1} \\ &= \{\{\text{Tr}_{A_1}[(M_{b_1|y_1}^{A_1} \otimes \mathbb{1}^{A_2})\rho^{A_1A_2}]\}_{b_1}\}_{y_1}, \end{aligned} \quad (36a)$$

$$\begin{aligned} \mathbf{M}_2^{A_2} \circ \mathbf{P}^{A_1A_2} &:= \{\{\tilde{\rho}_{b_2|y_2}^{A_1}\}_{b_2}\}_{y_2} \\ &= \{\{\text{Tr}_{A_2}[(\mathbb{1}^{A_1} \otimes M_{b_2|y_2}^{A_2})\rho^{A_1A_2}]\}_{b_2}\}_{y_2}. \end{aligned} \quad (36b)$$

Note that here and elsewhere in the paper, we will use $\tilde{\rho}$ to denote a subnormalized state. The operational identities satisfied by these multi-sources are of the form

$$\mathcal{O}(\mathbf{M}_1 \circ \mathbf{P}) = \{ \{ \alpha_{b_1 y_1} \} \mid \sum_{b_1 y_1} \alpha_{b_1 y_1} \tilde{\rho}_{b_1 | y_1}^{A_2} = \mathbb{0}^{A_2} \}, \quad (37a)$$

$$\mathcal{O}(\mathbf{M}_2 \circ \mathbf{P}) = \{ \{ \alpha_{b_2 y_2} \} \mid \sum_{b_2 y_2} \alpha_{b_2 y_2} \tilde{\rho}_{b_2 | y_2}^{A_1} = \mathbb{0}^{A_1} \}. \quad (37b)$$

Note that these last two sets of identities actually subsume those in Eqs. (35a) and (35b). To see that $\mathcal{O}(\mathbf{M}_1) \subseteq \mathcal{O}(\mathbf{M}_1 \circ \mathbf{P})$ for instance, it suffices to note that $\sum_{b_1 y_1} \alpha_{b_1 y_1} M_{b_1 | y_1}^{A_1} = \mathbb{0}^{A_1}$ implies $\sum_{b_1 y_1} \alpha_{b_1 y_1} \text{Tr}_1 [M_{b_1 | y_1}^{A_1} \otimes \mathbb{1}^{A_2} \rho^{A_1 A_2}] = \mathbb{0}^{A_2}$.

There is also a more subtle kind of operational identity analogous to the one in Eq. (13). Specifically, the unnormalized states $\tilde{\rho}_{b_2 | y_2}^{A_1}$ and $\tilde{\rho}_{b_1 | y_1}^{A_2}$ in the steering assemblage [i.e., in Eq. (36)] can also be composed in parallel, giving rise to bipartite states on $A_1 A_2$. Since such states are of the same type as the given bipartite state $\rho^{A_1 A_2}$, there might exist a new operational identity for some set of real coefficients $\{ \alpha_{b_1 b_2 y_1 y_2} \}_{b_1 b_2 y_1 y_2}$ such that

$$\rho^{A_1 A_2} = \sum_{b_1 b_2 y_1 y_2} \alpha_{b_1 b_2 y_1 y_2} \tilde{\rho}_{b_2 | y_2}^{A_1} \otimes \tilde{\rho}_{b_1 | y_1}^{A_2}. \quad (38)$$

Such an identity exists if and only if $\rho^{A_1 A_2}$ lies in the linear span of the set $\{ \tilde{\rho}_{b_2 | y_2}^{A_1} \otimes \tilde{\rho}_{b_1 | y_1}^{A_2} \}_{b_1 b_2 y_1 y_2}$. The operational identity of the form of Eq. (38) will be denoted by

$$\begin{aligned} \mathcal{O}(\tilde{\mathbf{P}}) &= \{ \{ \alpha_{b_1 b_2 y_1 y_2} \} \mid \rho^{A_1 A_2} \\ &= \sum_{b_1 b_2 y_1 y_2} \alpha_{b_1 b_2 y_1 y_2} \tilde{\rho}_{b_2 | y_2}^{A_1} \otimes \tilde{\rho}_{b_1 | y_1}^{A_2} \}. \end{aligned} \quad (39)$$

The circuit in Fig. 12(a) is classically explainable (i.e., admits an ontological model) if and only if the following conditions are met. One can introduce a separable ontic state space $\Lambda_1 \times \Lambda_2$ over which one can reproduce the quantum statistics via

$$p(b_1 b_2 | y_1 y_2) = \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) p(b_1 | y_1 \lambda_1) p(b_2 | y_2 \lambda_2), \quad (40)$$

where $p(b_1 | y_1 \lambda_1)$ satisfies the ontological identities specified by $\mathcal{O}(\mathbf{M}_1)$, and $p(b_2 | y_2 \lambda_2)$ satisfies the ontological identities specified by $\mathcal{O}(\mathbf{M}_2)$. Moreover, defining compositions

$$\tilde{p}(b_1 \lambda_2 | y_1) := \sum_{\lambda_1} p(b_1 | y_1 \lambda_1) p(\lambda_1 \lambda_2), \quad (41a)$$

$$\tilde{p}(b_2 \lambda_1 | y_2) := \sum_{\lambda_2} p(b_2 | y_2 \lambda_2) p(\lambda_1 \lambda_2), \quad (41b)$$

we require that $\tilde{p}(b_1 \lambda_2 | y_1)$ must satisfy the ontological identities specified by $\mathcal{O}(\mathbf{M}_1 \circ \mathbf{P})$, and $\tilde{p}(b_2 \lambda_1 | y_2)$ must satisfy the ontological identities specified by $\mathcal{O}(\mathbf{M}_2 \circ \mathbf{P})$. Finally, we require that $p(\lambda_1 \lambda_2)$ must satisfy the ontological identity specified by $\mathcal{O}(\tilde{\mathbf{P}})$. All told, we require that, for all λ_1, λ_2 ,

$$\sum_{b_1 y_1} \alpha_{b_1 y_1} p(b_1 | y_1 \lambda_1) = 0, \quad \forall \{ \alpha_{b_1 y_1} \} \in \mathcal{O}(\mathbf{M}_1), \quad (42a)$$

$$\sum_{b_2 y_2} \alpha_{b_2 y_2} p(b_2 | y_2 \lambda_2) = 0, \quad \forall \{ \alpha_{b_2 y_2} \} \in \mathcal{O}(\mathbf{M}_2), \quad (42b)$$

$$\sum_{b_1 y_1} \alpha_{b_1 y_1} \tilde{p}(b_1 \lambda_2 | y_1) = 0, \quad \forall \{ \alpha_{b_1 y_1} \} \in \mathcal{O}(\mathbf{M}_1 \circ \mathbf{P}), \quad (42c)$$

$$\sum_{b_2 y_2} \alpha_{b_2 y_2} \tilde{p}(b_2 \lambda_1 | y_2) = 0, \quad \forall \{ \alpha_{b_2 y_2} \} \in \mathcal{O}(\mathbf{M}_2 \circ \mathbf{P}), \quad (42d)$$

$$\begin{aligned} p(\lambda_1 \lambda_2) &= \sum_{b_1 b_2 y_1 y_2} \alpha_{b_1 b_2 y_1 y_2} \tilde{p}(b_2 \lambda_1 | y_2) \tilde{p}(b_1 \lambda_2 | y_1) \\ &\quad \forall \{ \alpha_{b_1 b_2 y_1 y_2} \} \in \mathcal{O}(\tilde{\mathbf{P}}). \end{aligned} \quad (42e)$$

Note that the first two constraints in Eq. (42) follow from linearity, while the latter three follow from diagram preservation together with linearity.

Importantly, since the ontological model for the steering assemblages has already been explicitly constructed in Eq. (41), this yields the following necessary condition for classical explainability.

Proposition 1. Any circuit with local measurements on a bipartite state [as depicted in Fig. 12(a)] is classically explainable only if the two prepare-measure circuits it induces by boxing processes together [in one or the other direction, as in Fig. 12(b)] are classically explainable.

The intuition behind this result is straightforward. To determine the classical explainability of a prepare-measure scenario obtained by boxing operation, it suffices to check only a subset of the constraints in Eq. (42) [specifically, only Eqs. (42b) and (42c) or only Eqs. (42a) and (42d)], rather than the full set that must be checked in the original bipartite scenario. Consequently, if a bipartite scenario is classically explainable, then so is each of the unipartite prepare-measure scenarios it defines by steering.

For completeness, a formal proof of this implication is provided in Appendix I.

B. The classical-nonclassical divide for bipartite states

We obtain the definition of classicality for a single bipartite state by particularizing Definition 3' to this type of process.

Definition 7. A state on a bipartite system is classical if and only if the statistics generated by the set of circuits where it is contracted with any product effect (i.e., where

the set ranges over all product effects) is classically explainable (in the sense of Definition 1).

We therefore consider a prepare-measure circuit of the type depicted in Fig. 12(a), with a single bipartite state and local measurements as depicted in Fig. 6(e), where the range of measurements includes every possible effect on A_1 and every possible effect on A_2 . A necessary condition for classicality of a bipartite state can then be obtained as a simple corollary of Proposition 1.

Proposition 2. A bipartite state $\rho^{A_1 A_2}$ is classical only if every multi-source on A_2 that is obtained by contracting $\rho^{A_1 A_2}$ with a multi-measurement on A_1 is classical, and every multi-source on A_1 that is obtained by contracting $\rho^{A_1 A_2}$ with a multi-measurement on A_2 is classical.

Proof. The proof is simple. If the bipartite state is classical, then by Definition 7, the circuit where it is contracted with any product effect is classically explainable. By Proposition 1, this implies that the two prepare-measure circuits [as depicted in Fig. 12(b)] are classically explainable when all local effects on each side are considered. But classical explainability for each prepare-measure circuit then implies that each corresponding set of steered states (the composition between the bipartite state and local effects on the other side) is classical according to Definition 5. ■

Whether the classicality of the set of steered states on both sides is sufficient to establish the classicality of the bipartite state is still unknown. However, in the following, we will provide the necessary and sufficient conditions for a classical bipartite state.

First, define the operator spaces

$$V_1 := \text{Span}(\{\text{Tr}_{A_2}[(\mathbb{1}^{A_1} \otimes M^{A_2})\rho^{A_1 A_2}]: M^{A_2} \in \mathcal{M}(\mathcal{H}^{A_2})\}), \quad (43a)$$

$$V_2 := \text{Span}(\{\text{Tr}_{A_1}[(M^{A_1} \otimes \mathbb{1}^{A_2})\rho^{A_1 A_2}]: M^{A_1} \in \mathcal{M}(\mathcal{H}^{A_1})\}). \quad (43b)$$

Whether or not a bipartite state $\rho^{A_1 A_2}$ is classical can then be characterized based on a product-frame decomposition, analogous to Theorems 2 and 3.

Theorem 4. A bipartite state $\rho^{A_1 A_2} \in \mathcal{S}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{A_2})$ is classical if and only if it admits a frame decomposition:

$$\rho^{A_1 A_2} = \sum_{\lambda_1, \lambda_2} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2})\rho^{A_1 A_2}] \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}, \quad (44)$$

where the frame on the operator space V_i is a fixed set of density operators $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i}$, and the dual frame is a set of Hermitian operators $\{H_{\lambda_i}^{A_i}\}_{\lambda_i}$ satisfying $\sum_{\lambda_i} H_{\lambda_i}^{A_i} = \mathcal{P}_{V_i}(\mathbb{1}^{A_i})$ with $\mathcal{P}_{V_i}$ denoting the superoperator projecting onto V_i , and such that

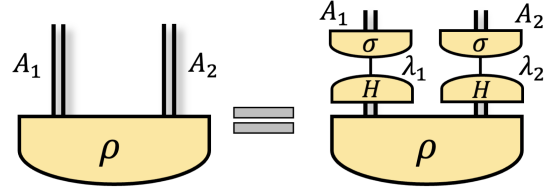


FIG. 13. A bipartite state is classical if and only if it admits a frame representation as shown on the right-hand side of the diagrammatic equation, where the condition on the frame and dual-frame operators are given in Theorem 4.

$$\text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2})\rho^{A_1 A_2}] \geq 0, \quad \forall \lambda_1, \lambda_2. \quad (45)$$

This condition is depicted in Fig. 13.

As a direct consequence of Theorem 4, it is easy to see that for a bipartite state, entanglement implies nonclassicality because satisfying Eqs. (44) and (45) in the condition for classicality articulated in Theorem 4 implies that $\rho^{A_1 A_2}$ must be separable.

Corollary 7. If a bipartite state is entangled, then it is nonclassical.

Within the set of all bipartite states, the boundary between separable and entangled is distinct from the boundary between steerable and unsteerable, and both of these are also distinct from the boundary between local and nonlocal [95–98]. As it turns out, the boundary between classical and nonclassical, as defined here, is also distinct from each of these. This follows from the fact that while separability is a necessary condition for classicality (the content of Corollary 7), it is not a sufficient condition, as is demonstrated by the following example.

Example 5. The separable bipartite state

$$\rho^{A_1 A_2} = \frac{1}{4}(|0\rangle\langle 0| \otimes |0\rangle\langle 0| + |1\rangle\langle 1| \otimes |1\rangle\langle 1| + |2\rangle\langle 2| \otimes |+\rangle\langle +| + |3\rangle\langle 3| \otimes |-\rangle\langle -|) \quad (46)$$

is nonclassical, because there exists a measurement on A_1 , namely, the projector-valued measure $\{|0\rangle\langle 0|, |1\rangle\langle 1|, |2\rangle\langle 2|, |3\rangle\langle 3|\}$ that steers A_2 to the “BB84 source” $\{\frac{1}{4}|0\rangle\langle 0|, \frac{1}{4}|1\rangle\langle 1|, \frac{1}{4}|+\rangle\langle +|, \frac{1}{4}|-\rangle\langle -|\}$, which was shown to be nonclassical in Example 2. The claim then follows from Proposition 2.

If a separable state can be nonclassical, it is natural to ask whether this nonclassicality constitutes a resource for achieving quantum advantages in any information-processing task. The answer is that it does (at least in some cases). Consider the separable state from Example 5 above. It induces the nonclassical BB84 source on A_2 , which is known to be a resource for quantum key distribution [91] and for parity-oblivious multiplexing [99]. That said, a complete understanding of the resourcefulness of separable states that are nonclassical—particularly whether

their resourcefulness extends beyond their ability to realize nonclassical resources via steering—remains open.

It is worth noting that, nonclassicality of a bipartite state, in the sense introduced here, does not coincide with having nonzero discord [100]; for example, the state $\frac{1}{2}|0\rangle\langle 0| \otimes |0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \otimes |+\rangle\langle +|$ has nonzero discord but is classical by our definition. It is not hard to see that nonzero discord is necessary but not sufficient for a single bipartite state to be nonclassical.

A complete characterization of the boundary between classical and nonclassical for bipartite states is likely to be difficult to obtain. For one, the set of classical bipartite states is not convex, because all product states are classical, but not all separable states are.

However, in the special case of a two-qubit system, a complete characterization is possible to obtain.

Proposition 3. A bipartite qubit state $\rho^{A_1A_2}$ is classical if and only if it is separable.

Proof. The “only if” direction follows from Corollary 7; we now prove the “if” direction.

We first treat the special case where $\rho^{A_2} := \text{Tr}_{A_1}[\rho^{A_1A_2}] = \frac{1}{2}$. By Ref. [101], for any two-qubit separable state with $\rho^{A_2} = \frac{1}{2}$, the set of states that can be steered on A_1 (i.e., the steering ellipsoid) is contained in a tetrahedron in the Bloch ball. Let d denote the dimension of the steering ellipsoid of $\rho^{A_1A_2}$ (an integer less than or equal to 3), Ref. [101] shows constructively that $\rho^{A_1A_2}$ admits a decomposition

$$\rho^{A_1A_2} = \sum_{\lambda_1, \lambda_2=0}^d p_{\lambda_1} \delta_{\lambda_1, \lambda_2} \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}, \quad (47)$$

where $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i=0}^d$ are linearly independent sets of states. Given the linear independence, there exist unique dual frames $\{H_{\lambda_i}^{A_i}\}_{\lambda_i=0}^d$ on the space $V_i := \text{Span}(\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i=0}^d)$ such that $\text{Tr}[H_{\lambda_i}^{A_i} \sigma_{\lambda'_i}^{A_i}] = \delta_{\lambda_i, \lambda'_i}$ [one can infer that $\sum_{\lambda_i} H_{\lambda_i}^{A_i} = \mathcal{P}_{V_i}(\mathbb{1})$]. Therefore, one can construct the following frame decomposition for $\rho^{A_1A_2}$:

$$\rho^{A_1A_2} = \sum_{\lambda_1, \lambda_2=0}^d \text{Tr}[H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2} \rho^{A_1A_2}] \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}, \quad (48)$$

where $\text{Tr}[H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2} \rho^{A_1A_2}] = p_{\lambda_1} \delta_{\lambda_1, \lambda_2} \geq 0$. We have shown that the conditions of Theorem 4 are satisfied, and so it follows that any two-qubit separable state $\rho^{A_1A_2}$ satisfying $\text{Tr}_{A_1}[\rho^{A_1A_2}] = \frac{1}{2}$ is classical.

Next, consider the case where $\rho^{A_2} \neq \frac{1}{2}$. If ρ^{A_2} is rank-1, $\rho^{A_1A_2}$ must be of the form $\rho^{A_1A_2} = \rho^{A_1} \otimes \rho^{A_2}$ and thus is classical. If ρ^{A_2} has full rank, we can obtain a locally filtered state $\tilde{\rho}^{A_1A_2}$ via

$$\tilde{\rho}^{A_1A_2} := [\mathbb{1}^{A_1} \otimes (2\rho^{A_2})^{-\frac{1}{2}}] \rho^{A_1A_2} [\mathbb{1}^{A_1} \otimes (2\rho^{A_2})^{-\frac{1}{2}}], \quad (49)$$

so that $\text{Tr}_{A_1}[\tilde{\rho}^{A_1A_2}] = \frac{1}{2}$. Since this local filtering preserves separability, $\tilde{\rho}^{A_1A_2}$ is still separable. Now, given dual-frame operators $\{\tilde{\sigma}_{\lambda_i}^{A_i}\}_{\lambda_i}$ and $\{\tilde{H}_{\lambda_i}^{A_i}\}_{\lambda_i}$ for $\tilde{\rho}^{A_1A_2}$ defined on the operator space $\tilde{V}_i := \text{Span}(\{\tilde{\sigma}_{\lambda_i}^{A_i}\}_{\lambda_i=0}^d)$, one can map these frame operators back to $\rho^{A_1A_2}$ by defining

$$\sigma_{\lambda_1}^{A_1} := \tilde{\sigma}_{\lambda_1}^{A_1}, \quad H_{\lambda_1}^{A_1} = \tilde{H}_{\lambda_1}^{A_1}, \quad (50a)$$

$$\sigma_{\lambda_2}^{A_2} := \frac{(2\rho^{A_2})^{\frac{1}{2}} \tilde{\sigma}_{\lambda_2}^{A_2} (2\rho^{A_2})^{\frac{1}{2}}}{\text{Tr}[\tilde{\sigma}_{\lambda_2}^{A_2} (2\rho^{A_2})]}, \quad (50b)$$

$$H_{\lambda_2}^{A_2} := \text{Tr}[\tilde{\sigma}_{\lambda_2}^{A_2} (2\rho^{A_2})] \left[(2\rho^{A_2})^{-\frac{1}{2}} \tilde{H}_{\lambda_2}^{A_2} (2\rho^{A_2})^{-\frac{1}{2}} \right]. \quad (50c)$$

Then, one has the frame representation

$$\rho^{A_1A_2} = \sum_{\lambda_1, \lambda_2=0}^d \text{Tr}[H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2} \rho^{A_1A_2}] \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}. \quad (51)$$

It is easy to see that $\text{Tr}[H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2} \rho^{A_1A_2}] \geq 0$, and moreover, $\{H_{\lambda_1}^{A_1}\}$ and $\{\sigma_{\lambda_1}^{A_1}\}$ define a dual basis on V_1 , and $\sum_{\lambda_1} H_{\lambda_1}^{A_1} = \mathcal{P}_{V_1}(\mathbb{1})$ since they are untouched. In order to show that $\sum_{\lambda_2} H_{\lambda_2}^{A_2} = \mathcal{P}_{V_2}(\mathbb{1})$, it suffices to note that from Eqs. (50b) and (50c), we have

$$\begin{aligned} \text{Tr}[H_{\lambda_2}^{A_2} \sigma_{\lambda'_2}^{A_2}] &= \frac{\text{Tr}[\tilde{\sigma}_{\lambda_2}^{A_2} (2\rho^{A_2})]}{\text{Tr}[\tilde{\sigma}_{\lambda'_2}^{A_2} (2\rho^{A_2})]} \text{Tr}[\tilde{H}_{\lambda_2}^{A_2} \tilde{\sigma}_{\lambda'_2}^{A_2}] \\ &= \frac{\text{Tr}[\tilde{\sigma}_{\lambda_2}^{A_2} (2\rho^{A_2})]}{\text{Tr}[\tilde{\sigma}_{\lambda'_2}^{A_2} (2\rho^{A_2})]} \delta_{\lambda_2, \lambda'_2} = \delta_{\lambda_2, \lambda'_2}, \end{aligned} \quad (52)$$

where the second equality holds because $\{\sigma_{\lambda_2}^{A_2}\}_{\lambda_2}$ and $\{\tilde{H}_{\lambda_2}^{A_2}\}_{\lambda_2}$ are a pair of basis and dual basis.

Again, having shown that the conditions of Theorem 4 are satisfied, it follows that the two-qubit state $\rho^{A_1A_2}$ is classical. ■

Remark. The proof can be run in the opposite direction, i.e., by considering the steering ellipsoid on A_2 .

Using an analogous proof technique, we can find a sufficient condition for classicality for bipartite states, independent of the subsystem dimensions.

Corollary 8. A bipartite state $\rho^{A_1A_2}$ is classical if it admits a separable decomposition of the form

$$\rho^{A_1A_2} = \sum_{\lambda_1, \lambda_2} p(\lambda_1, \lambda_2) \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}, \quad (53)$$

where each of $\{\sigma_{\lambda_1}^{A_1}\}_{\lambda_1}$ and $\{\sigma_{\lambda_2}^{A_2}\}_{\lambda_2}$ is a set of linearly independent states.

For subsystems having Hilbert space dimensionality greater than two, there exist separable states that do not admit such a decomposition, as is illustrated in Example 5. This fact allows the possibility of separable nonclassical states when one moves beyond qubits.

In the following, we briefly discuss a conjecture concerning the task of witnessing the nonclassicality of a bipartite state. It is based on a property we term nonsingularity (which we will define right after the conjecture).

Conjecture 1. A nonsingular bipartite state is classical if and only if it is separable.

The nonsingularity property is defined as follows.

Definition 8. A bipartite state $\rho^{A_1 A_2}$ is said to be nonsingular if the steering maps induced by $\rho^{A_1 A_2}$ are invertible. The steering maps are the superoperators Ψ_1^ρ and Ψ_2^ρ defined by the expressions

$$\Psi_1^\rho(X) = \text{Tr}_{A_1}[(X^T \otimes \mathbb{1})\rho^{A_1 A_2}], \quad (54)$$

$$\Psi_2^\rho(X) = \text{Tr}_{A_2}[(\mathbb{1} \otimes X^T)\rho^{A_1 A_2}], \quad (55)$$

where T denotes the transpose. An equivalent characterization of what it is for a bipartite state to be nonsingular can be given in terms of operational identities: A bipartite state is nonsingular if and only if the coefficients defining the linear identities holding among the measurement effects on A_1 are equivalent to those defining the linear identities holding among the steered states on A_2 , and vice versa; that is, if and only if $\mathcal{O}(\mathbf{M}_1 \circ \mathbf{P}) = \mathcal{O}(\mathbf{M}_1)$ and $\mathcal{O}(\mathbf{M}_2 \circ \mathbf{P}) = \mathcal{O}(\mathbf{M}_2)$ [recall that $\mathcal{O}(\mathbf{M}_1)$, $\mathcal{O}(\mathbf{M}_2)$, $\mathcal{O}(\mathbf{M}_1 \circ \mathbf{P})$, and $\mathcal{O}(\mathbf{M}_2 \circ \mathbf{P})$ are defined in Eqs. (35) and (37) respectively]. One can summarize this characterization of nonsingularity as follows: A bipartite state is nonsingular if it does not induce any new operational identities when contracted with the local effects.

That this characterization of nonsingularity coincides with that of Definition 8 can be easily seen from the fact that $\mathcal{O}(\mathbf{M}_i \circ \mathbf{P})$ in Eq. (37) can be expressed as

$$\mathcal{O}(\mathbf{M}_i \circ \mathbf{P}) = \{ \{ \alpha_{b_i y_i} \} \mid \sum_{b_i y_i} \alpha_{b_i y_i} \Psi_i^\rho([M_{b_i y_i}^{A_i}]^T) = 0 \}, \quad (56)$$

but using the invertibility of Ψ_i^ρ and the invertibility of the transpose, this implies

$$\mathcal{O}(\mathbf{M}_i \circ \mathbf{P}) = \{ \{ \alpha_{b_i y_i} \} \mid \sum_{b_i y_i} \alpha_{b_i y_i} M_{b_i y_i}^{A_i} = 0 \}, \quad (57)$$

such that, recalling the definition of $\mathcal{O}(\mathbf{M}_i)$ from Eqs. (35a) and (35b), we infer that $\mathcal{O}(\mathbf{M}_i \circ \mathbf{P}) = \mathcal{O}(\mathbf{M}_i)$.

Corollary 7 establishes that for bipartite states, being separable is a necessary condition for being classical.

To establish Conjecture 1, what remains to be proven is that for nonsingular bipartite states, being separable is also a sufficient condition for being classical. When considering what constraints the operational identities imply for a non-contextual ontological model, one needs only to consider the first two sets of constraints in Eq. (42), which are of the form

$$\sum_{b_i y_i} \alpha_{b_i y_i} p(b_i | y_i, \lambda_i) = 0, \quad \forall \{ \alpha_{b_i y_i} \} \in \mathcal{O}(\mathbf{M}_i). \quad (58)$$

This is because the other two sets of constraints in Eqs. (42c) and (42d), namely,

$$\sum_{b_i y_i} \alpha_{b_i y_i} \sum_{\lambda_j} p(\lambda_i \lambda_j) p(b_i | y_i, \lambda_i) = 0 \quad \forall \{ \alpha_{b_i y_i} \} \in \mathcal{O}(\mathbf{M}_i \circ \mathbf{P}), \quad (59)$$

are automatically satisfied since we have $\mathcal{O}(\mathbf{M}_i \circ \mathbf{P}) = \mathcal{O}(\mathbf{M}_i)$ for nonsingular bipartite states.

Now, let us discuss why Conjecture 1 might hold and why we are unable to provide a complete proof. Consider a bipartite state; from the proof of Corollary 7, the state is separable if and only if the statistics generated by the set of circuits where it is contracted with any product effect admits an ontological model that satisfies the first two sets of identities in Eq. (42). If the bipartite state is also nonsingular, as we discussed above, we can disregard the third and fourth sets of constraints. In other words, statistics generated by a nonsingular bipartite state automatically satisfy the first four sets of identities in Eq. (42).

What remains unclear is whether the fifth constraint in Eq. (42)—namely, the one implied by the operational identity of $\mathcal{O}(\tilde{\mathbf{P}})$ —has nontrivial consequences for the characterization of classicality for bipartite states. Thus, it is still an open question whether every nonsingular separable state is classical.

As discussed at the beginning of this section, diagram preservation implies that an ontological model assigns a Cartesian product of ontic state spaces to a particular tensor product of Hilbert spaces. This leads to additional constraints on the ontological representation of quantum processes—constraints that can in turn imply nonclassicality of a process that would otherwise be judged classical. This highlights how conclusions about nonclassicality hold only relative to a choice of factorization of the Hilbert space. For instance, as noted above, a four-dimensional Hilbert space can be factorized into a pair of two-dimensional Hilbert spaces in many ways, where the factor spaces correspond to virtual subsystems [102]. Because for a given pure state there is always some choice of factorization relative to which it is entangled, it follows from Corollary 7 that there is always some choice of factorization relative to which it is nonclassical. In this article, the

choice of factorization of Hilbert space relative to which nonclassicality is being judged is always specified by the circuit one is considering.

Note that the ideas and results in this section generalize straightforwardly to multipartite states, and that some of these generalizations appear in Sec. VII.

VI. PROCESSES OF THE TRANSFORMATION VARIETY

To discuss the nonclassicality of a process of the transformation variety (a channel, multi-channel, instrument, or multi-instrument), one must consider the circuit obtained by contracting the process with its dual process, namely, a comb, as depicted in Fig. 6(d). But given Theorem 1, we need not consider all combs. It suffices to consider those that factorize, meaning they can be understood as a preparation preceding the process and a measurement following it. In other words, it suffices to consider circuits that have the form depicted in Fig. 14(a), termed a prepare-transform-measure scenario.

As in the previous sections, we begin by reviewing the form of a classical explanation (i.e., a diagram-preserving and linear ontological representation) of a prepare-transform-measure scenario. We consider the case where the preparation and measurement stages are general (a multi-state and a multi-measurement, respectively), but the transformation is simply a channel, rather than a general process of the transformation variety. We label the quantum system before and after the transformation by A and B .

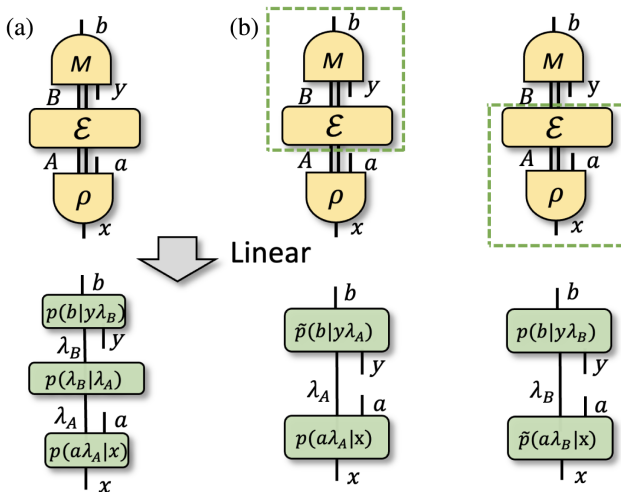


FIG. 14. (a) A prepare-transform-measure circuit. (b) By boxing the channel together with the preparation or the measurement, we can obtain a pair of prepare-measure circuits. At the top of the figure, we depict the quantum circuits; at the bottom, we depict an ontological model of each circuit.

A. Classical explainability of a prepare-transform-measure scenario

We consider a multi-source $\mathbf{P}^A := \{ \{ p(a|x) \rho_{a|x}^A \}_a \}_x$ on $\mathcal{H}^A$, followed by a channel $\mathbf{T}^{B|A} := \mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$, and a multi-measurement $\mathbf{M}^B := \{ \{ M_{b|y}^B \}_b \}_y$ on $\mathcal{H}^B$. For such a circuit, quantum theory predicts that

$$p(ab|xy) = p(a|x) \text{Tr} [M_{b|y}^B \mathcal{E}^{B|A}(\rho_{a|x}^A)]. \quad (60)$$

In this circuit, there are operational identities among the states on the input system and among the effects on the output system, namely,

$$\mathcal{O}(\mathbf{M}) = \{ \{ \alpha_{by} \} | \sum_{by} \alpha_{by} M_{b|y}^B = \mathbb{0}^B \}, \quad (61a)$$

$$\mathcal{O}(\mathbf{P}) = \{ \{ \alpha_{ax} \} | \sum_{ax} \alpha_{ax} p(a|x) \rho_{a|x}^A = \mathbb{0}^A \}. \quad (61b)$$

There are also additional operational identities, as a consequence of the fact that the sequential composition of the channel, together with the multi-source on its input, defines an effective multi-source, and the sequential composition of the channel, together with the multi-measurement on its output, defines an effective multi-measurement. Specifically, the effective multi-source and multi-measurement are

$$\mathbf{M}^B \circ \mathbf{T}^{B|A} := \{ \{ \tilde{M}_{b|y}^A \}_b \}_y = \{ \{ [\mathcal{E}^{B|A}]^\dagger(M_{b|y}^B) \}_b \}_y, \quad (62a)$$

$$\mathbf{T}^{B|A} \circ \mathbf{P}^A := \{ \{ \tilde{\rho}_{a|x}^B \}_a \}_x = \{ \{ p(a|x) \mathcal{E}^{B|A}(\rho_{a|x}^A) \}_a \}_x, \quad (62b)$$

and the operational identities for these are of the form

$$\mathcal{O}(\mathbf{M} \circ \mathbf{T}) = \{ \{ \alpha_{by} \} | \sum_{by} \alpha_{by} \tilde{M}_{b|y}^A = \mathbb{0}^A \}, \quad (63a)$$

$$\mathcal{O}(\mathbf{T} \circ \mathbf{P}) = \{ \{ \alpha_{ax} \} | \sum_{ax} \alpha_{ax} \tilde{\rho}_{a|x}^B = \mathbb{0}^B \}. \quad (63b)$$

Moreover, the sequential composition of the new multi-measurement on system A , $\{ \{ \tilde{M}_{b|y}^A \}_b \}_y$, with the new multi-source on system B , $\{ \{ \tilde{\rho}_{a|x}^B \}_a \}_x$, gives rise to a measure-and-prepare multi-instrument, namely,

$$\tilde{\mathcal{E}}_{ab|xy}^{B|A}(\cdot) := \tilde{\rho}_{a|x}^B \text{Tr} [\tilde{M}_{b|y}^A(\cdot)]. \quad (64)$$

Since this composite process is of the same type as the given channel $\mathcal{E}^{B|A}$ making up the prepare-transform-measure scenario, there might exist a new operational identity in the form of

$$\mathcal{E}^{B|A} = \sum_{abxy} \alpha_{abxy} \tilde{\mathcal{E}}_{ab|xy}^{B|A}. \quad (65)$$

Such an identity exists if and only if $\mathcal{E}^{B|A}$ lies in the linear span of the set $\{\tilde{\mathcal{E}}_{ab|xy}^{B|A}(\cdot) = \tilde{\rho}_{a|x}^B \text{Tr}[\tilde{M}_{b|y}^A(\cdot)]\}_{abxy}$ with $\tilde{\rho}_{a|x}^B$ and $\tilde{M}_{b|y}^A$ given in Eq. (62). The operational identity of the form of Eq. (65) will then be denoted by

$$\mathcal{O}(\tilde{\mathcal{T}}) = \{\{\alpha_{abxy}\} | \mathcal{E}^{B|A} = \sum_{abxy} \alpha_{abxy} \tilde{\mathcal{E}}_{ab|xy}^{B|A}\}. \quad (66)$$

Consequently, a prepare-transform-measure circuit is classically explainable if and only if one can reproduce the quantum statistics via

$$p(ab|xy) = \sum_{\lambda_A \lambda_B} p(b|y\lambda_B) p(\lambda_B|\lambda_A) p(a\lambda_A|x), \quad (67)$$

where $p(b|y\lambda_B)$ respects the identities implied by $\mathcal{O}(\mathbf{M})$, and $p(a\lambda_A|x)$ respects the identities implied by $\mathcal{O}(\mathbf{P})$. Moreover, define

$$\tilde{p}(b|y\lambda_A) := \sum_{\lambda_B} p(b|y\lambda_B) p(\lambda_B|\lambda_A), \quad (68a)$$

$$\tilde{p}(a\lambda_B|x) := \sum_{\lambda_A} p(a\lambda_A|x) p(\lambda_B|\lambda_A). \quad (68b)$$

Here, $\tilde{p}(b|y\lambda_A)$ must satisfy the ontological identities implied by $\mathcal{O}(\mathbf{M} \circ \mathbf{T})$, and $\tilde{p}(a\lambda_B|x)$ must satisfy the ontological identities implied by $\mathcal{O}(\mathbf{T} \circ \mathbf{P})$, while $p(\lambda_B|\lambda_A)$ must satisfy the ontological identities implied by $\mathcal{O}(\tilde{\mathcal{T}})$. Explicitly, we require that for all λ_A, λ_B ,

$$\sum_{by} \alpha_{by} p(b|y\lambda_B) = 0, \quad \forall \{\alpha_{by}\} \in \mathcal{O}(\mathbf{M}), \quad (69a)$$

$$\sum_{ax} \alpha_{ax} p(a\lambda_A|x) = 0, \quad \forall \{\alpha_{ax}\} \in \mathcal{O}(\mathbf{P}), \quad (69b)$$

$$\sum_{by} \alpha_{by} \tilde{p}(b|y\lambda_A) = 0, \quad \forall \{\alpha_{by}\} \in \mathcal{O}(\mathbf{M} \circ \mathbf{T}), \quad (69c)$$

$$\sum_{ax} \alpha_{ax} \tilde{p}(a\lambda_B|x) = 0, \quad \forall \{\alpha_{ax}\} \in \mathcal{O}(\mathbf{T} \circ \mathbf{P}), \quad (69d)$$

$$p(\lambda_B|\lambda_A) = \sum_{abxy} \alpha_{abxy} \tilde{p}(b|y\lambda_A) \tilde{p}(a\lambda_B|x) \quad \forall \{\alpha_{abxy}\} \in \mathcal{O}(\tilde{\mathcal{T}}). \quad (69e)$$

The first two constraints here follow from linearity, while the latter three follow from diagram preservation together with linearity. Finally, we present a result that will be useful

for the characterization of nonclassical channels (the analog of Proposition 1).

Proposition 4. Any prepare-transform-measure circuit [as depicted in Fig. 14(a)] is classically explainable only if the two prepare-measure circuits it induces by boxing the transformation with the preparation or with the measurements—the two cases depicted in Fig. 14(b)—are classically explainable.

B. The classical-nonclassical divide for channels

We turn now to the definition of classicality for a channel, again obtained by particularizing Definition 3'. Recalling that the most general dual process of a channel from A to B is a comb with output A and input B and that factorizing such a comb constitutes a preparation on A and a measurement on B , we are led to consider a prepare-transform-measure circuit of the type depicted in Fig. 14(a), but where the range of preparations includes every possible state on A and the range of measurements includes every possible effect on B —a situation depicted in Fig. 6(f).

Definition 9. A channel is classical if and only if the statistics generated by the set of circuits where it is contracted with any state at its input and any effect at its output (i.e., where the set ranges over all states and all effects) is classically explainable (in the sense of Definition 1).

A useful result concerning the sufficient condition for nonclassical channels (necessary condition for classical channels) can be inferred from Proposition 4.

Proposition 5. A channel $\mathcal{E}^{B|A}$ is classical only if the image of every multi-source on A under $\mathcal{E}^{B|A}$ is a classical multi-source on B , and the image of every multi-measurement on B under $[\mathcal{E}^{B|A}]^\dagger$ is a classical multi-measurement on A .

It is straightforward to give a proof that is analogous to that of Proposition 2. The result can also be understood as following from Proposition 2 by leveraging Lemma 1 (which will be presented in Sec. VII) because the latter result shows that the verdict regarding the nonclassicality of a channel is the same as that regarding the bipartite state that is Choi isomorphic to it.

However, the converse of Proposition 5 is yet unknown. Instead, we show that whether a quantum channel $\mathcal{E}^{B|A}$ is classical or not can be characterized by the following theorem, analogous to Theorems 2–4, which is proven in Appendix J 2.

Define the operator spaces

$$V_B = \text{Span}[\{\mathcal{E}^{B|A}(\rho^A) : \rho^A \in \mathcal{S}(\mathcal{H}^A)\}], \quad (70a)$$

$$V_A = \text{Span}[\{[\mathcal{E}^{B|A}]^\dagger(M^B) : M^B \in \mathcal{M}(\mathcal{H}^B)\}]. \quad (70b)$$

Note that since $[\mathcal{E}^{B|A}]^\dagger$ is a unital channel, V_A must also contain the identity operator $\mathbb{1}^A$.

Theorem 5. A channel $\mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$ is classical if and only if it admits a frame representation:

$$\mathcal{E}^{B|A}(\cdot) = \sum_{\lambda_A \lambda_B} \sigma_{\lambda_B}^B \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A(\cdot)], \quad (71)$$

where the frame on V_A is a set of Hermitian operators $\{F_{\lambda_A}^A\}_{\lambda_A}$ each with unit trace $\text{Tr}[F_{\lambda_A}^A] = 1$, and the dual frame on V_A is a POVM $\{G_{\lambda_A}^A\}_{\lambda_A}$ satisfying $\sum_{\lambda_A} G_{\lambda_A}^A = \mathcal{P}_{V_A}(\mathbb{1}^A) = \mathbb{1}^A$, while the frame on V_B is a set of density operators $\{\sigma_{\lambda_B}^B\}_{\lambda_B}$, and the dual frame on V_B is a set of Hermitian operators $\{H_{\lambda_B}^B\}_{\lambda_B}$ satisfying $\sum_{\lambda_B} H_{\lambda_B}^B = \mathcal{P}_{V_B}(\mathbb{1}^B)$, with $\mathcal{P}_{V_{B(A)}}$ denoting the superoperator projecting onto $V_{B(A)}$, and such that

$$\text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \geq 0, \quad \forall \lambda_A, \lambda_B. \quad (72)$$

This condition is depicted in Fig. 15.

As a direct consequence, one can observe that a quantum channel being classical implies it is entanglement breaking.

Corollary 9. If a channel is non-entanglement-breaking, then it is nonclassical.

This also follows immediately from the fact that if a state of a bipartite system is entangled, then it is nonclassical (Corollary 7) together with the fact that nonclassicality of processes is preserved under the Choi isomorphism (Lemma 1, which will be presented further on in the article).

Although being entanglement breaking is a necessary condition for classicality, it is not sufficient, as illustrated by the following examples.

Example 6. The entanglement-breaking channel

$$\mathcal{E}(X) = \text{Tr}[|0\rangle\langle 0|X]|0\rangle\langle 0| + \text{Tr}[|1\rangle\langle 1|X]|1\rangle\langle 1| + \text{Tr}[|2\rangle\langle 2|X]|+\rangle\langle +| + \text{Tr}[|3\rangle\langle 3|X]|-\rangle\langle -| \quad (73)$$

is nonclassical. This follows from Proposition 5 and the fact that if the input to the channel is the source $\{\frac{1}{4}|i\rangle\langle i|\}_{i=0}^3$, then its output is the source $\{\frac{1}{4}|0\rangle\langle 0|, \frac{1}{4}|1\rangle\langle 1|, \frac{1}{4}|+\rangle\langle +|, \frac{1}{4}|-\rangle\langle -|\}$, which is the nonclassical source discussed in

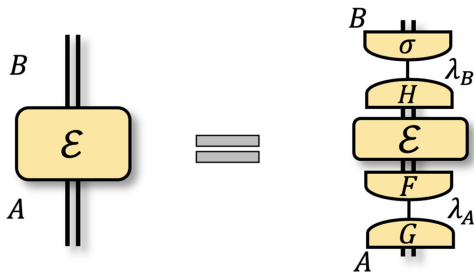


FIG. 15. A quantum channel is classical if and only if it admits a frame representation as shown on the right-hand side of the diagrammatic equation, where the condition on the frame and dual-frame operators are given in Theorem 5.

Example 2. (Note that there does not exist a set of effects that is mapped by the adjoint of this channel to a non-classical set of effects, since any set of effects will get mapped to a set of effects that are all diagonal in the same eigenbasis, namely, $\{|i\rangle\langle i|\}_{i=0}^3$.)

Example 7. The entanglement-breaking channel

$$\mathcal{E}(X) = \frac{1}{2} [\text{Tr}[|0\rangle\langle 0|X]|0\rangle\langle 0| + \text{Tr}[|1\rangle\langle 1|X]|1\rangle\langle 1| + \text{Tr}[|+\rangle\langle +|X]|2\rangle\langle 2| + \text{Tr}[|-\rangle\langle -|X]|3\rangle\langle 3|] \quad (74)$$

is also nonclassical. This follows from Proposition 5 and the fact that if the output to the channel is composed with the POVM $\{|i\rangle\langle i|\}_{i=0}^3$, then it defines an effective POVM $\{\frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|1\rangle\langle 1|, \frac{1}{2}|+\rangle\langle +|, \frac{1}{2}|-\rangle\langle -|\}$, which is nonclassical according to our Corollary 6.

An alternative proof of the nonclassicality of this channel is to note that the channel is Choi isomorphic to the nonclassical separable state in Example 5 and to leverage the fact that nonclassicality is preserved under the Choi isomorphism (Lemma 1).

VII. PROCESSES OF ARBITRARY TYPE

In this section, we derive results on the nonclassicality of a general process that extends those in the previous sections. We first present some necessary conditions for classicality for the process of arbitrary type. Next, we present a full characterization of such a classicality in terms of a frame representation. Finally, we show that the nonclassicality of any process can be deduced from the nonclassicality of an associated multipartite state.

Here, the most general process we consider is a multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$, which has a classical setting variable z and a classical outcome variable c , takes a set of quantum inputs labeled $\vec{A} := (A_1, \dots, A_n)$ and produces a set of quantum outputs labeled $\vec{B} := (B_1, \dots, B_m)$. This class subsumes multi-source, multi-measurement, bipartite, multi-instrument, and multipartite versions of each. We also refer the reader to Appendix K for the generalization to N -combs [103].

This section introduces substantial notation to treat processes of arbitrary type. Conceptually, however, the results are very similar to those in Secs. V and VIB for bipartite states and channels. Readers who wish to avoid the notational overhead may focus on the main results—Theorem 6, Proposition 6, Proposition 7, Corollary 10, and Corollary 11 and consult the surrounding material only as needed.

A. Classical explainability of a multipartite prepare-transform-measure scenario

As preparation for our discussion of the nonclassicality of a general multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$,

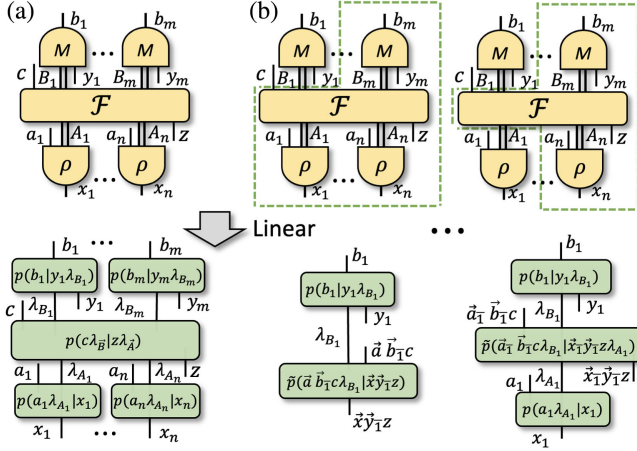


FIG. 16. (a) A multipartite prepare-transform-measure circuit with local multi-measurements, local multi-sources, and a multipartite multi-instrument. (b) By boxing processes together, we can, for instance, view this circuit as a unipartite prepare-measure circuit on B_1 , or as a unipartite prepare-transform-measure circuit from A_1 to B_1 , where we denote $\vec{b}_1 := \{b_2, \dots, b_m\}$ (and similarly for $\vec{a}_1, \vec{x}_1$, and $\vec{y}_1$). At the top of the circuit, we depict the quantum circuit, and at the bottom, we depict their diagram-preserving ontological model.

we first consider the question of when a multipartite prepare-transform-measure circuit involving a multi-instrument, along with local preparations and local measurements on the multipartite system, is classically explainable.

Specifically, we consider the circuit depicted in Fig. 16 where at the preparation stage, there is for each $i \in [n] := \{1, \dots, n\}$, a multi-source $P_i^{A_i} := \{\{p(a_i|x_i)\rho_{a_i|x_i}^{A_i}\}_{a_i}\}_{x_i}$ on $\mathcal{H}^{A_i}$, at the transformation stage there is a multipartite multi-instrument $T^{\vec{B}|\vec{A}} := \{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ where $\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}: \mathcal{L}(\otimes_{i \in [n]} \mathcal{H}^{A_i}) \mapsto \mathcal{L}(\otimes_{j \in [m]} \mathcal{H}^{B_j})$, and finally, at the measurement stage there is, for each $j \in [m] := \{1, \dots, m\}$, a local measurement $M_j^{B_j} := \{\{M_{b_j|y_j}^{B_j}\}_{b_j}\}_{y_j}$. For such a circuit, quantum theory predicts the statistics

$$p(\vec{a} \vec{b} c | \vec{x} \vec{y} z) = \text{Tr} \left[\left(\bigotimes_{j \in [m]} M_{b_j|y_j}^{B_j} \right) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\bigotimes_{i \in [n]} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right) \right]. \quad (75)$$

Similar to earlier discussions of simpler quantum circuits, one can enumerate all operational identities among the individual processes in the circuit. Specifically,

$$\mathcal{O}(P_i) = \{\{\alpha_{a_i x_i}\} | \sum_{a_i} \alpha_{a_i x_i} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} = \mathbb{O}^{A_i}\}, \quad (76a)$$

$$\mathcal{O}(M_j) = \{\{\alpha_{b_j y_j}\} | \sum_{b_j} \alpha_{b_j y_j} M_{b_j|y_j}^{B_j} = \mathbb{O}^{B_j}\}, \quad (76b)$$

$$\mathcal{O}(T) = \{\{\alpha_{cz}\} | \sum_{cz} \alpha_{cz} \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} = \mathbb{O}^{\vec{B}|\vec{A}}\}. \quad (76c)$$

There are additional operational identities in this circuit arising from sequential compositions of the multipartite multi-instrument with multi-states on different subsets of the input systems and with multi-measurements on different subsets of the output systems, respectively.

Specifically, for any sets $\mathcal{I} \subseteq [n]$ and $\mathcal{J} \subseteq [m]$ with complementary sets $\bar{\mathcal{I}} := [n]/\mathcal{I}$ and $\bar{\mathcal{J}} := [m]/\mathcal{J}$, we denote $\vec{B}_{\mathcal{J}} := (B_j)_{j \in \mathcal{J}}$ and $\vec{B}_{\bar{\mathcal{J}}} := (B_j)_{j \in \bar{\mathcal{J}}}$ with analogous notation for $\vec{A}_{\mathcal{I}}, \vec{A}_{\bar{\mathcal{I}}}, \vec{b}_{\mathcal{J}}$, and $\vec{x}_{\mathcal{I}}$. We can define the effective process (a transformation from $\vec{A}_{\bar{\mathcal{I}}}$ to $\vec{B}_{\mathcal{J}}$) by

$$\begin{aligned} T^{\vec{B}_{\mathcal{J}}|\vec{A}_{\bar{\mathcal{I}}}}(\cdot) &= \bigotimes_{j \in \mathcal{J}} M_j^{B_j} \circ T^{\vec{B}|\vec{A}} \circ \bigotimes_{i \in \bar{\mathcal{I}}} P_i^{A_i} \\ &:= \{\{\tilde{\mathcal{F}}_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c | \vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z}^{\vec{B}_{\mathcal{J}}|\vec{A}_{\bar{\mathcal{I}}}}(\cdot)\}_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c}\}_{\vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z} \\ &= \{\{\text{Tr}_{\vec{B}_{\mathcal{J}}} [M_{\vec{b}_{\mathcal{J}}|\vec{y}_{\mathcal{J}}}^{\vec{B}} \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} (\tilde{\rho}_{\vec{a}_{\mathcal{I}}|\vec{x}_{\mathcal{I}}}^{\vec{A}_{\bar{\mathcal{I}}}} \otimes (\cdot))] \}_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c}\}_{\vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z}, \end{aligned} \quad (77)$$

where we denote $M_{\vec{b}_{\mathcal{J}}|\vec{y}_{\mathcal{J}}}^{\vec{B}} := \bigotimes_{j' \in \bar{\mathcal{J}}} \mathbb{1}^{B_{j'}} \otimes \bigotimes_{j \in \mathcal{J}} M_{b_j|y_j}^{B_j}$ and $\tilde{\rho}_{\vec{a}_{\mathcal{I}}|\vec{x}_{\mathcal{I}}}^{\vec{A}_{\bar{\mathcal{I}}}} := \bigotimes_{i \in \mathcal{I}} p(a_i|x_i) \rho_{a_i|x_i}^{A_i}$. For these effective processes, their operational identities are of the form

$$\mathcal{O}(T^{\vec{B}_{\mathcal{J}}|\vec{A}_{\bar{\mathcal{I}}}}) = \left\{ \{\alpha_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c}\} | \sum_{\substack{\vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z \\ \vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c}} \alpha_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c} \tilde{\mathcal{F}}_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c | \vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z}^{\vec{B}_{\mathcal{J}}|\vec{A}_{\bar{\mathcal{I}}}} = \mathbb{O}^{\vec{B}_{\mathcal{J}}|\vec{A}_{\bar{\mathcal{I}}}} \right\}, \quad (78)$$

of which the operational identities in Eq. (76c) are a special case with $\mathcal{I} = \mathcal{J} = \emptyset$.

In addition to these sequential compositions, parallel compositions of the effective processes in Eq. (77) can also lead to nontrivial operational identities. For instance, given any disjoint collection of subsets of $[m]$, $\{\bar{\mathcal{J}}_{\nu}\}_{\nu}$, and any disjoint collection of subsets of $[n]$, $\{\bar{\mathcal{I}}_{\nu}\}_{\nu}$, such that $\cup_{\nu} \bar{\mathcal{J}}_{\nu} = [m]$ and $\cup_{\nu} \bar{\mathcal{I}}_{\nu} = [n]$. Denote $\mathcal{I}_{\nu} := [n]/\bar{\mathcal{I}}_{\nu}$ and $\mathcal{J}_{\nu} := [m]/\bar{\mathcal{J}}_{\nu}$ as the corresponding set-theoretic complements, there might exist nontrivial operational identities of the form

$$\begin{aligned} \sum_{cz} \alpha_{cz} \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \\ + \sum_{\substack{\{\vec{x}_{\mathcal{I}_{\nu}}^{\nu} \vec{y}_{\mathcal{J}_{\nu}}^{\nu} z^{\nu}\}_{\nu} \\ \{\vec{a}_{\mathcal{I}_{\nu}}^{\nu} \vec{b}_{\mathcal{J}_{\nu}}^{\nu} c^{\nu}\}_{\nu} \\ \{\vec{a}_{\mathcal{I}_{\nu}}^{\nu} \vec{b}_{\mathcal{J}_{\nu}}^{\nu} c^{\nu}\}_{\nu}}} \alpha_{\{\vec{a}_{\mathcal{I}_{\nu}}^{\nu} \vec{b}_{\mathcal{J}_{\nu}}^{\nu} c^{\nu}\}_{\nu}} \bigotimes_{\nu} \tilde{\mathcal{F}}_{\vec{a}_{\mathcal{I}_{\nu}}^{\nu} \vec{b}_{\mathcal{J}_{\nu}}^{\nu} c^{\nu} | \vec{x}_{\mathcal{I}_{\nu}}^{\nu} \vec{y}_{\mathcal{J}_{\nu}}^{\nu} z^{\nu}}^{\vec{B}_{\mathcal{J}_{\nu}}|\vec{A}_{\mathcal{I}_{\nu}}}} = \mathbb{O}^{\vec{B}|\vec{A}}, \end{aligned} \quad (79)$$

where the superscript ν in c^{ν} , $\vec{a}^{\nu}$ is an index to distinguish different settings and outcomes in different parts of the

composition. One can think of this kind of operational identity as a generalization of Eqs. (38) and (65).

Since the new identity obviously includes the one in Eq. (76c) as a special case, the sets of new operational identities are

$$\begin{aligned} \mathcal{O}(\tilde{\mathbf{T}}) := & \left\{ \{\alpha_{cz}\} \cup \{\alpha_{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu}}\} \mid \sum_{cz} \alpha_{cz} \mathcal{F}_{c|z}^{\tilde{B}|\tilde{A}} \right. \\ & \left. + \sum_{\substack{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu} \\ \{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} c^{\nu}\}_{\nu}}} \alpha_{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu}} \otimes \tilde{\mathcal{F}}_{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} c^{\nu} | \tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}}^{\tilde{B}|\tilde{A}} = 0^{\tilde{B}|\tilde{A}} \right\}, \end{aligned} \quad (80)$$

where the coefficients $\{\alpha_{cz}\}$ and $\{\alpha_{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu}}\}$ are not both uniformly zero.

Consequently, the circuit in Fig. 16 is classically explainable if and only if one can reproduce the quantum statistics via

$$\begin{aligned} p(\vec{a} \vec{b} c | \vec{x} \vec{y} z) &= \sum_{\lambda_{\tilde{B}} \lambda_{\tilde{A}}} p(c \lambda_{\tilde{B}} | z \lambda_{\tilde{A}}) \prod_j p(b_j | y_j \lambda_{B_j}) \\ &\quad \times \prod_i p(a_i \lambda_{A_i} | x_i), \end{aligned} \quad (81)$$

where $p(c \lambda_{\tilde{B}} | z \lambda_{\tilde{A}})$, $\{p(b_j | y_j \lambda_{B_j})\}_j$, and $\{p(a_i \lambda_{A_i} | x_i)\}_i$ satisfy all the relevant ontological identities. Moreover, for any choice of $\mathcal{I} \subseteq [n]$ and $\mathcal{J} \subseteq [m]$, if we can define the effective ontological processes associated with $\mathbf{T}_{\tilde{\mathcal{I}}\tilde{\mathcal{J}}}^{\tilde{B}|\tilde{A}}$ as

$$\begin{aligned} p(\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c \lambda_{\tilde{B}_{\mathcal{J}}} | \vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z \lambda_{\tilde{A}_{\mathcal{I}}}) \\ = \sum_{\lambda_{\tilde{B}} \lambda_{\tilde{A}}} p(c \lambda_{\tilde{B}} | z \lambda_{\tilde{A}}) \prod_{j \in \mathcal{J}} p(b_j | y_j \lambda_{B_j}) \prod_{i \in \mathcal{I}} p(a_i \lambda_{A_i} | x_i), \end{aligned} \quad (82)$$

then they must satisfy the following ontological identities for all $\lambda_{\tilde{B}}, \lambda_{\tilde{A}}$:

$$\sum_{b_j y_j} \alpha_{b_j y_j} p(b_j | y_j \lambda_{B_j}) = 0, \quad (83a)$$

$$\sum_{a_i x_i} \alpha_{a_i x_i} p(a_i \lambda_{A_i} | x_i) = 0, \quad (83b)$$

$$\sum_{\substack{\vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z \\ \vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c}} \alpha_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c} p(\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c \lambda_{\tilde{B}_{\mathcal{J}}} | \vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z \lambda_{\tilde{A}_{\mathcal{I}}}) = 0, \quad (83c)$$

$$\begin{aligned} \sum_{\substack{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu} \\ \{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} c^{\nu}\}_{\nu}}} \alpha_{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu}} \prod_{\nu} p(\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} c^{\nu} \lambda_{\tilde{B}_{\mathcal{J}}} | \tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu} \lambda_{\tilde{A}_{\mathcal{I}}}) \\ + \sum_{cz} \alpha_{cz} p(c \lambda_{\tilde{B}} | z \lambda_{\tilde{A}}) = 0 \end{aligned} \quad (83d)$$

for all $\{\alpha_{b_j y_j}\} \in \mathcal{O}(\mathbf{M}_j)$, $\{\alpha_{a_i x_i}\} \in \mathcal{O}(\mathbf{P}_i)$, $\{\alpha_{\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} c}^{\vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} z}\} \in \mathcal{O}(\mathbf{T}_{\tilde{\mathcal{I}}\tilde{\mathcal{J}}}^{\tilde{B}|\tilde{A}})$, and $\{\alpha_{cz}\} \cup \{\alpha_{\{\tilde{a}_{\mathcal{I}\nu}^{\nu} \tilde{b}_{\mathcal{J}\nu}^{\nu} z^{\nu}\}_{\nu}}\} \in \mathcal{O}(\tilde{\mathbf{T}})$.

It is, in general, very difficult to check all of these identities and decide if the circuit is classically explainable; however, the following proposition gives a necessary condition for it to be classically explainable, which is similar to the ones in Propositions 1 and 4.

Proposition 6. Any multipartite prepare-transform-measure circuit [of the type depicted in Fig. 16(a)] is classically explainable only if any circuit it induces by boxing the transformation with any subset of local preparations and/or local measurements [as depicted in Fig. 16(b)] is classically explainable.

B. The classical-nonclassical divide for a general process

One can always box the multipartite multi-instrument together with local preparations on all input subsystems and local measurements on all but one of the output subsystems, as depicted in Fig. 16(b), to obtain an effective preparation on the remaining output subsystem. Similarly, one can always box the multipartite multi-instrument with local measurements on all output subsystems and local preparations on all but one of the input subsystems to obtain an effective measurement on the remaining input subsystem. In this way, a multipartite prepare-transform-measure scenario defines a number of effective prepare-measure scenarios. Applying Proposition 6, we obtain the following necessary condition for nonclassicality of a multipartite multi-instrument.

Proposition 7. A multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\tilde{B}|\tilde{A}}\}_c\}_z$ is classical only if every multi-source on quantum output B_j that is obtained by contracting it with a multi-measurement on each of the remaining quantum outputs and a multi-source on each of the quantum inputs is classical, and every multi-measurement on quantum input A_i that is obtained by contracting it with a multi-source on each of the remaining quantum inputs and a multi-measurement on each of the quantum outputs is classical.

The proof is similar to that of Proposition 2. If the multipartite multi-instrument is classical, then the circuit where it is contracted with the full set of product effects and the full set of product states [depicted in Fig. 16(a)] is classically explainable. By Proposition 6, this implies that all prepare-measure circuits induced by boxing the prepare-transform-measure circuit [e.g., Fig. 16(b)] are classically explainable, which indicates that the multi-states and multi-measurements that appear in the prepare-measure circuit are classical in the sense of Definitions 5 and 6. However, whether the converse is true is yet unknown.

Similar to Theorems 2–5, we here provide the necessary and sufficient characterization of classicality for multipartite multi-instruments, which is given by the following theorem.

Define V_{B_j} as the operator space spanned by all possible states on $\mathcal{H}^{B_j}$ that can be generated by $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ when it is boxed with local preparations on all subsystems $\{A_i\}_{i=1}^n$ and local measurements on all subsystems but subsystem B_j . Specifically,

$$V_{B_j} := \text{Span}(\{\text{Tr}_{\vec{B}_j}[(\mathbb{1}^{B_j} \otimes M^{\vec{B}_j})\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\rho^{\vec{A}})]\}_{c,z}), \quad (84)$$

with $\rho^{\vec{A}} = \otimes_{i \in [n]} \rho^{A_i}$ ranging over all $\rho^{A_i} \in \mathcal{S}(\mathcal{H}^{A_i})$, and $M^{\vec{B}_j} = \otimes_{j' \neq j} M^{B_{j'}}$ ranging over all $M^{B_{j'}} \in \mathcal{M}(\mathcal{H}^{B_{j'}})$.

Similarly, define V_{A_i} as the operator space spanned by all possible effects on $\mathcal{H}^{A_i}$ that can be generated by $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ when it is boxed with local measurements on all subsystems $\{B_j\}_{j=1}^m$ and local preparations on all subsystems but A_i . Specifically,

$$V_{A_i} := \text{Span}(\{\text{Tr}_{\vec{A}_i}[(\mathbb{1}^{A_i} \otimes \rho^{\vec{A}_i})[\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}]^\dagger(M^{\vec{B}})]\}_{c,z}), \quad (85)$$

with $\rho^{\vec{A}_i} = \otimes_{i' \neq i} \rho^{A_{i'}}$ ranging over all $\rho^{A_{i'}} \in \mathcal{S}(\mathcal{H}^{A_{i'}})$, and $M^{\vec{B}} = \otimes_{j \in [m]} M^{B_j}$ ranging over all $M^{B_j} \in \mathcal{M}(\mathcal{H}^{B_j})$.

Theorem 6. A multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$, where $\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}: \mathcal{L}(\otimes_{i=1}^n \mathcal{H}^{A_i}) \mapsto \mathcal{L}(\otimes_{j=1}^m \mathcal{H}^{B_j})$, is classical if and only if it admits a frame representation

$$\begin{aligned} \mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\cdot) &= \sum_{\vec{\lambda} \in \vec{\Lambda}} \otimes_j \sigma_{\lambda_{B_j}}^{B_j} \text{Tr}[(\otimes_j H_{\lambda_{B_j}}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\otimes_i F_{\lambda_{A_i}}^{A_i})] \\ &\quad \times \text{Tr}[(\otimes_i G_{\lambda_{A_i}}^{A_i})(\cdot)], \end{aligned} \quad (86)$$

where the frame on the operator space V_{B_j} is a fixed set of density operators $\{\sigma_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$, and the dual frame is a set of Hermitian operators $\{H_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$ satisfying $\sum_{\lambda_{B_j}} H_{\lambda_{B_j}}^{B_j} = \mathcal{P}_{V_{B_j}}(\mathbb{1}^{B_j})$; the frame on the operator space V_{A_i} is a set of Hermitian operators $\{F_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ satisfying $\text{Tr}[F_{\lambda_{A_i}}^{A_i}] = 1$, and the dual frame is a POVM $\{G_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ satisfying $\sum_{\lambda_{A_i}} G_{\lambda_{A_i}}^{A_i} = \mathcal{P}_{V_{A_i}}(\mathbb{1}^{A_i}) = \mathbb{1}^{A_i}$, with $\mathcal{P}_{V_{B_j}(A_i)}$ denoting the superoperator projecting onto $V_{B_j(A_i)}$, and such that

$$\text{Tr}[(\otimes_j H_{\lambda_{B_j}}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\otimes_i F_{\lambda_{A_i}}^{A_i})] \geq 0, \quad \forall \vec{\lambda}, c, z. \quad (87)$$

Such a frame decomposition is depicted diagrammatically in Fig. 17. The proof of this theorem is similar to that of Theorems 4 and 5, and is outlined in detail in Appendix J3 for completeness.

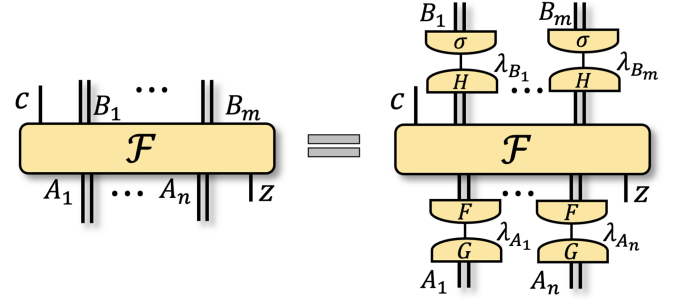


FIG. 17. A multipartite multi-instrument is classical if and only if it admits a frame representation as shown on the right-hand side of the equation, where the condition on the frames and dual frames are given in Theorem 6.

As a direct corollary of Theorem 6, we obtain the following necessary condition for the classicality (sufficient condition for the nonclassicality) of a multipartite state.

Corollary 10. A multipartite state is classical only if it is separable.

Moreover, as another direct corollary of Theorem 6, another necessary condition for the classicality of a multipartite state can be found in a manner analogous to Proposition 7.

Corollary 11. A multipartite state $\rho^{A_1, \dots, A_N}$ is classical only if, for every party A_i , any multi-sources on A_i obtained by contracting the multipartite state $\rho^{A_1, \dots, A_N}$ with any multi-measurements on the other parties $A_{i'} \neq A_i$ is classical.

In the multipartite case (just as in the bipartite case), separability is necessary but not sufficient for classicality. That is, there exist multipartite states that are not entangled across any partition but that are nonetheless nonclassical according to our proposal. It suffices to consider multipartite generalizations of Example 5.

One can also find tripartite qubit separable states whose steering assemblages are nonclassical on every party (recall that by contrast, as shown in Proposition 3, any bipartite qubit separable state is classical).

Example 8. Given a tripartite state defined as

$$\rho^{A_1 A_2 A_3} = \frac{1}{4} (|000\rangle\langle 000| + |-+1\rangle\langle -+1| \quad (88)$$

$$+ |1-+\rangle\langle 1-+| + |+1-\rangle\langle +1-|), \quad (89)$$

there exist measurements on any two subsystems that steer the remaining system to a set of states that is nonclassical. In particular, if any two parties each perform the BB84 measurement, which is associated with the POVM $\{\frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|1\rangle\langle 1|, \frac{1}{2}|+\rangle\langle +|, \frac{1}{2}|-\rangle\langle -|\}$, then they steer the third party's system to the ensemble containing the unnormalized states

$$\begin{aligned} & \left\{ \frac{3}{64} \left(\frac{2}{3} \rho_x + \frac{1}{3} \tau_y \right), \frac{3}{64} \left(\frac{1}{3} \rho_x + \frac{2}{3} \tau_y \right) \right\}_{x,y} \\ & \cup \left\{ \frac{5}{64} \left(\frac{3}{5} \rho_x + \frac{1}{5} \mathbb{1} \right), \frac{5}{64} \left(\frac{1}{5} \rho_x + \frac{2}{5} \mathbb{1} \right) \right\}_x \\ & \cup \left\{ \frac{5}{64} \left(\frac{3}{5} \tau_y + \frac{1}{5} \mathbb{1} \right), \frac{5}{64} \left(\frac{1}{5} \tau_y + \frac{2}{5} \mathbb{1} \right) \right\}_y, \end{aligned} \quad (90)$$

where x and y are binary variables with $\{\rho_x\}_x = \{|0\rangle\langle 0|, |1\rangle\langle 1|\}$ and $\{\tau_y\}_y = \{|+\rangle\langle +|, |-\rangle\langle -|\}$. This ensemble can be shown to be nonclassical using the semidefinite program given in Ref. [72].

The tripartite state above is the uniform mixture of the four product states that have been shown to exhibit “non-locality without entanglement” [104]. These product states form an unextendible product basis and cannot be perfectly distinguished by measurements implementable using local operations and classical communication (LOCC).

An interesting open question is whether, when assessing the nonclassicality of a multipartite state, it suffices to consider the nonclassicality of the multi-sources to which one can steer on each subsystem of the full partition. We have not managed to settle this question one way or the other. If considering single-party steering possibilities does suffice, this will greatly simplify the task of characterizing nonclassicality of multipartite states (and hence, general processes). If it does not, then this will constitute a new way in which the nonclassicality of a state can be judged to be genuinely multipartite [4,105–107].

C. Nonclassicality of a process via the nonclassicality of an associated multipartite state

Now, we will demonstrate how the nonclassicality of a multipartite multi-instrument can be deduced by analyzing the nonclassicality of an associated multipartite state.

Theorem 7. For any quantum process, define its unique associated multipartite state by

- (i) applying the Choi isomorphism to all of its quantum inputs,
- (ii) applying flag convexification to all of its classical inputs (setting variables),
- (iii) reconceptualizing all of its classical outputs (outcomes) as dephased quantum outputs.

The process is nonclassical if and only if its associated multipartite state is nonclassical.

This theorem follows from three lemmas.

The first lemma states that verdicts concerning the nonclassicality of a given process is preserved under the Choi isomorphism. [This is analogous to how any resource in the resource theory of local operations and shared randomness (LOSR) can always be transformed via LOSR operations into a process with only classical outputs and quantum inputs in such a way that the property of being nonfree is preserved [3].]

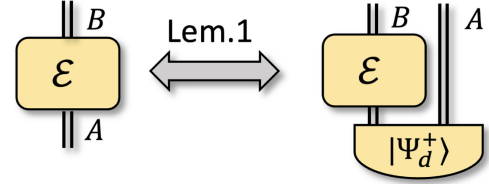


FIG. 18. As an example of Lemma 1, a channel $\mathcal{E}^{B|A}$ is nonclassical if and only if the Choi-isomorphic bipartite state $\rho_{\mathcal{E}}^{BA} = (\mathcal{E}^{B|A} \otimes \mathbb{1}^A) |\Psi^+\rangle\langle\Psi^+|^{AA}$ is nonclassical. (Here, the second A denotes the isomorphic copy used in the Choi construction).

Lemma 1. The state that is Choi isomorphic to a given process is nonclassical if and only if the process itself is nonclassical.

This is proven in Appendix L.

Hence, to decide nonclassicality, one can first map any process to one with no quantum inputs using the Choi isomorphism. This is depicted in Fig. 18. This connection enables, e.g., characterizations of nonclassicality for bipartite states to be transferred to channels and those for unipartite sources to be transferred to unipartite measurements.

The second lemma states that the verdict about the nonclassicality of a given process is preserved under mapping a classical input (a setting variable) into a classical output (an outcome variable) via flag convexification, a notion introduced in Ref. [31]. In flag convexification, the classical input variable is sampled from a full-support probability distribution, then copied, and the copy becomes a classical output variable. We depict an example of this on the right-hand side of Fig. 19. (In this example and henceforth, we take the probability distribution in question to be the uniform distribution, although our results hold just as well for any other full-support distribution.)

Lemma 2. The image of a given process under flag convexification of its classical input is nonclassical if and only if the process itself is nonclassical.

The proof follows from facts about the ontological representation of flag convexification described in Refs. [31,66]. Specifically, it follows from the fact that all that changes when one moves from a given process to its

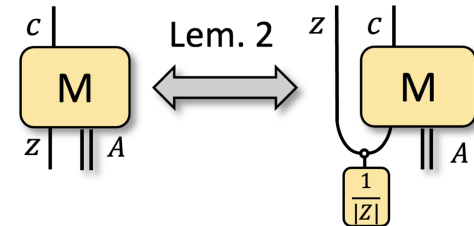


FIG. 19. As an example of Lemma 2, a multi-measurement $\{\{M_{c|z}\}_c\}_z$ is nonclassical if and only if the single measurement $\{\tilde{M}_{c,z}\}_{c,z}$ obtained from flag convexification of the setting variable z is nonclassical. Assuming the uniform distribution for sampling, we have $\tilde{M}_{c,z} = (1/|Z|)M_{c|z}$.

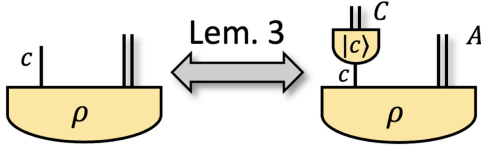


FIG. 20. As an example of Lemma 3, a source $\{p(c)\rho_c^A\}_c$ is nonclassical if and only if the bipartite state $\tilde{\rho}^{CA} = \sum_c p(c)|c\rangle\langle c|^C \otimes \rho_c^A$ obtained by treating its classical outcome as a dephased quantum output, is nonclassical. Here, the dephased quantum output is represented by preparing one of the orthogonal set of states $\{|c\rangle\}_c$ on system C conditional on the classical outcome c .

flag-convexified counterpart is that it is rescaled by a constant factor, and so an ontological representation of the given process can simply be rescaled by the same constant factor to give an ontological representation of the flag-convexified process. This lemma also follows from the fact that the frame-representation characterization of the classicality of a process (as in Theorem 6) is preserved under flag convexification of the setting variable.

The third lemma states that the nonclassicality of a process does not depend on whether or not one reconceptualizes its classical output (its outcome) as a quantum output that is dephased in a fixed basis.

Lemma 3. Whether or not a process is nonclassical (in the sense of Definition 3') is independent of whether one reconceptualizes its classical outputs as dephased quantum outputs.

This is proven in Appendix L and an example is depicted in Fig. 20.

(In the next section, we show that the analogous result is also true for classical inputs.)

Theorem 7 follows immediately by combining these three lemmas as depicted in Fig. 21. This implies that characterizations of nonclassicality of arbitrary multipartite states can be transferred to the characterization of the nonclassicality of arbitrary processes in quantum theory.

VIII. CONSISTENCY UNDER RECONCEPTUALIZING SETTINGS AND OUTCOMES AS DEPHASED QUANTUM SYSTEMS

A basic consistency condition for any notion of classicality is that it should not depend on the theorist's representational choices. In particular, it should not depend on whether one represents the settings and outcomes of a given process in the conventional manner, as classical variables, or as quantum systems that are dephased in an appropriate basis. This is an instance of the principle of cut motility which is introduced and defined in Ref. [108] as follows: For any physical theory purporting to have universal applicability, describing an experimental phenomenon typically involves making a cut between (i) the systems that are modeled explicitly within the theory and

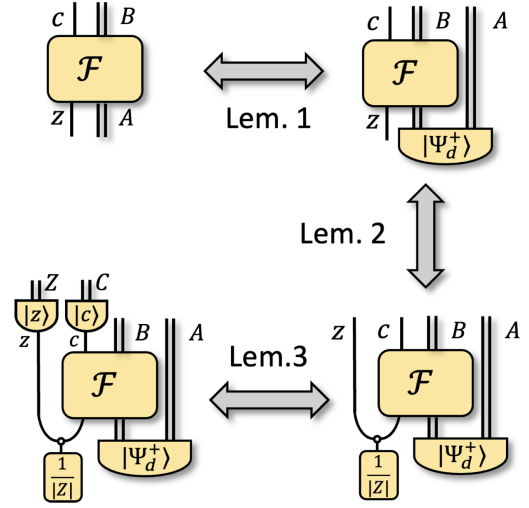


FIG. 21. A multi-instrument $\{\{\mathcal{F}_{c|z}^{B|A}\}_c\}_z$ is nonclassical if and only if the associated multipartite state $\rho^{ZCBA} = \sum_{cz} (1/|Z|) p(c|z) |z\rangle\langle z|^Z \otimes |c\rangle\langle c|^C \otimes \rho_{c|z}^{BA}$ is nonclassical, where $p(c|z)\rho_{c|z}^{BA} = (\mathcal{F}_{c|z}^{B|A} \otimes \mathbb{1}^A) |\Psi^+\rangle\langle\Psi^+|_{AA}$.

are the objects of interventions (preparations, measurements, transformations), and (ii) the systems that make up the devices that implement these interventions. A theory is said to satisfy the principle of cut motility if its predictions are independent of the placement of the cut.

The notion of classicality we have introduced here satisfies this consistency condition. In the case of reconceptualization of classical outputs, Lemma 3 already gives the result, and we can use the three lemmas from the previous section to extend this result to reconceptualization of classical settings, following the logic shown in Fig. 22.

Although this kind of consistency may seem quite basic, naïve attempts to stitch together different notions of classicality found in the literature can fail to be consistent in this manner and so can fail to provide a unified notion of classicality.

Consider a multi-state on a bipartite system, as shown in the left-hand side of Fig. 23. If one endorses the view that entanglement is necessary for nonclassicality of a bipartite state (call this the nonclassical-only-if-entangled view), it is reasonable to endorse the view that the nonclassical bipartite multi-states are all and only those for which there is at least one value of the setting variable that yields an entangled state. Now consider any bipartite multi-state for which the output system C is dephased in some basis. Taking the nonclassical-only-if-entangled view would imply that such a “one-output-dephased” process must be considered classical.

But the type of process just described can also be conceptualized as a multi-source on a unipartite system. Specifically, the dephased quantum system C can be reconceptualized as the classical outcome c of a multi-source on A . This is depicted on the right-hand side

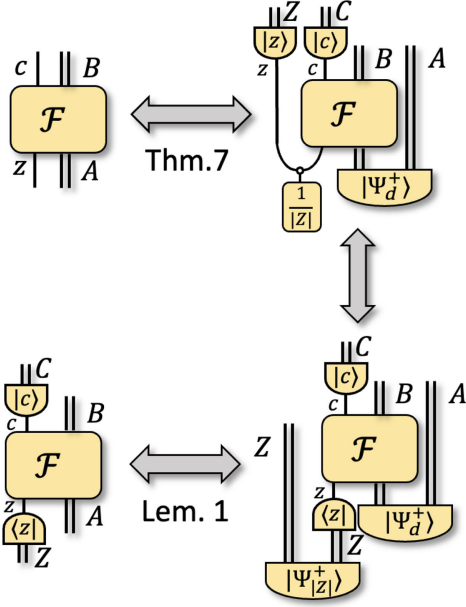


FIG. 22. The proof that nonclassicality is preserved under reconceptualizing classical inputs and outputs as dephased quantum outputs. The implication described by the vertical arrow follows from the identity $\text{Tr}_2[(1 \otimes |z\rangle\langle z|)|\Psi_{|z|}^+\rangle\langle\Psi_{|z|}^+|] = (1/|Z|)|z\rangle\langle z|$, where $|\Psi_{|z|}^+\rangle = (1/\sqrt{|Z|}) \sum_{z=1}^{|Z|} |z\rangle|z\rangle$.

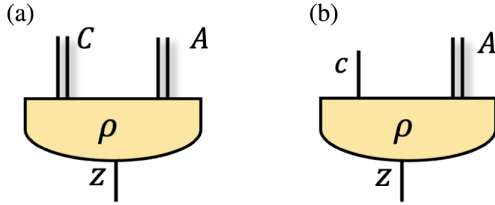


FIG. 23. Example of a challenge to the consistency criterion when reconceptualizing a classical outcome as a dephased quantum output. (a) A multi-state on a bipartite system; (b) a multi-source on a unipartite system. We can obtain an instance of the former from an instance of the latter by reconceptualizing the classical outcome as a dephased quantum output. Every bipartite multi-state obtained in this way contains only separable states, such that if one adopted the conventional view that entanglement is necessary for nonclassicality, it would be deemed classical. Meanwhile, if the unipartite multi-source is of the steerable variety, then under the conventional view that steerability is sufficient for nonclassicality, the same process would be judged to be nonclassical.

of Fig. 23. Now recall that there is a conventional notion of nonclassicality for a special class of multi-sources (those obtained in a steering scenario), namely, steerability [84]. Consequently, some processes of this type (namely, the steerable ones, or equivalently, the LOSR-nonfree ones [3,88]) will be judged to be nonclassical under this view.

In this way, the consistency criterion may be violated. Note that regardless of whether anyone has ever endorsed

this pair of type-specific notions of nonclassicality, they appear very reasonable at first glance and yet they fail the consistency test. A concrete example of this type is as follows.

Example 9. Consider the following multi-source on a unipartite system (i.e., assemblage):

$$\left\{ \left\{ \frac{1}{2}|0\rangle\langle 0|, \frac{1}{2}|1\rangle\langle 1| \right\}, \left\{ \frac{1}{2}|+\rangle\langle +|, \frac{1}{2}|-\rangle\langle -| \right\} \right\}.$$

Note that it is a steerable assemblage. If one treats the classical outcome variable as a dephased quantum output, the process is reconceptualized as a multi-state on a bipartite system, namely, the one associated with the set of bipartite states $\{\rho_1, \rho_2\}$, where

$$\rho_1 = \frac{1}{2}|0\rangle\langle 0| \otimes |0\rangle\langle 0| + \frac{1}{2}|1\rangle\langle 1| \otimes |1\rangle\langle 1|, \quad (91)$$

$$\rho_2 = \frac{1}{2}|0\rangle\langle 0| \otimes |+\rangle\langle +| + \frac{1}{2}|1\rangle\langle 1| \otimes |-\rangle\langle -|. \quad (92)$$

Note that both of these states are separable. If one takes the view that the classical-nonclassical divide for a multi-source is the unsteerable-steerable divide, then one will judge the process to be nonclassical. Meanwhile, if one takes the view that a multi-state on a bipartite system is nonclassical only if it yields an entangled state for at least one value of the setting variable, then one will judge the process to be classical [109].

Example 9 therefore illustrates how seemingly reasonable views about the classical-nonclassical divide for different types of processes need not satisfy our consistency criterion. It follows that achieving such a consistency is not trivial, and that if one is proposing a unified notion of nonclassicality, one should check for such a consistency. (It is worth noting that the notion of resourcefulness in the LOSR resource theory is indeed consistent in this sense [3].)

For the unified notion of nonclassicality proposed in this paper, the process in Example 9 above is found to be nonclassical regardless of how it is conceptualized. Indeed, the requirement of consistency in this example provides further evidence that containing only separable states is not a sufficient condition for a bipartite multi-state to be considered classical.

A second example of how seemingly reasonable type-specific notions of nonclassicality can fail to satisfy the consistency criterion is as follows.

Consider a multi-channel, as shown in the left-hand side of Fig. 24. If one holds the view that for bipartite states, being entangled is necessary for nonclassicality, then it is also reasonable to take the view that for channels, being non-entanglement-breaking is necessary for nonclassicality. Furthermore, it is reasonable to then endorse the following view of nonclassicality for multi-channels:

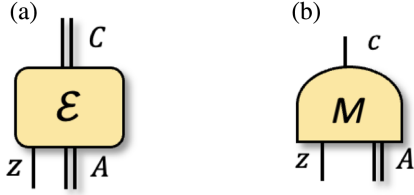


FIG. 24. Another example of a challenge to the consistency criterion when reconceptualizing a classical outcome as a dephased quantum output. (a) A multi-channel; (b) a multi-measurement. We can obtain an instance of the former from an instance of the latter by reconceptualizing the classical outcome as a dephased quantum output. Every multi-channel obtained in this way contains only entanglement-breaking channels, such that if one adopted the conventional view that for channels, being non-entanglement-breaking is necessary for nonclassicality, the multi-channel would be deemed classical. Meanwhile, if the multi-measurement contains incompatible measurements, then, under the conventional view that such an incompatibility is sufficient for nonclassicality, the same process would be judged to be nonclassical.

The nonclassical multi-channels are all and only those for which the set of channels accessible by variation of the setting variable contains at least one that is non-entanglement-breaking.

Now consider any multi-channel from A to C for which the output system C is dephased in some basis. Under the view just described, such a process must be considered classical.

But the type of process just described can also be conceptualized as a multi-measurement on a unipartite system. Specifically, the dephased quantum system C can be reconceptualized as the classical outcome c of a multi-measurement on A . This is depicted on the right-hand side of Fig. 24.

Now recall that a conventional view about the classical-nonclassical divide for multi-measurements (equivalently, sets of measurements) is that incompatibility is necessary for nonclassicality. Consequently, some of the multi-measurements that correspond to a multi-channel with a dephased output will be judged to be nonclassical under this view.

In this way, the consistency criterion may be violated. The following provides a concrete example.

Example 10. Consider the following multi-measurement (equivalently, set of measurements):

$$\{|0\rangle\langle 0|, |1\rangle\langle 1|, \{|+\rangle\langle +|, |-\rangle\langle -|\}\}.$$

Note that the two measurements in this set are incompatible. If one treats the classical outcome variable as a dephased quantum output, the process is reconceptualized as a multi-channel, namely, the one associated with the set of channels $\{\mathcal{E}_1, \mathcal{E}_2\}$, where

$$\mathcal{E}_1(\cdot) = \text{Tr}[|0\rangle\langle 0|(\cdot)]|0\rangle\langle 0| + \text{Tr}[|1\rangle\langle 1|(\cdot)]|1\rangle\langle 1|, \quad (93)$$

$$\mathcal{E}_2(\cdot) = \text{Tr}[|+\rangle\langle +|(\cdot)]|0\rangle\langle 0| + \text{Tr}[|-\rangle\langle -|(\cdot)]|1\rangle\langle 1|. \quad (94)$$

Note that both of these channels are entanglement breaking.

If one takes the view that the classical-nonclassical divide for a multi-measurement is the compatible-incompatible divide, then one will judge the process to be nonclassical. Meanwhile, if one takes the view that a multi-channel is nonclassical only if it yields a non-entanglement-breaking channel for at least one value of the setting variable, then one will judge the process to be classical.

Again, we see that seemingly reasonable views about the classical-nonclassical divide for different types of processes need not satisfy our consistency criterion.

For the unified notion of nonclassicality proposed here, the process in Example 10 is found to be nonclassical regardless of how it is conceptualized. This possibility follows from the fact that, relative to our notion, incompatibility in a multi-measurement is a sufficient condition for nonclassicality, and a multi-channel can be deemed nonclassical even if all the channels it contains are entanglement breaking [110].

IX. RELATED WORKS

Another approach to defining a classical-nonclassical divide in a manner that follows a unified principle which applies across a wide range of types of processes is the type-independent resource theory presented in Refs. [3,4]. This approach defines a notion of nonclassicality of common cause, which applies to processes such as multipartite states, multipartite boxes (i.e., processes with only classical inputs and classical outputs), assemblages, and so on. However, the classical-nonclassical divide defined by this approach is trivial for unipartite processes, and it is moreover not clear how to generalize this approach to arbitrary causal structures (although first steps were discussed in Appendix A.3 of Ref. [2]).

It is nonetheless useful to contrast this approach with the one we propose here. One way to understand the relationship between the two is that classicality, as defined in the LOSR approach, demands that the ontological accounts of the statistics have a causal structure that mirrors that of the quantum circuit (an instance of the assumption of diagram preservation), while the approach described here demands, in addition, that the ontological representation map be linear. This stronger assumption is motivated by the principle of generalized noncontextuality and implies that the set of processes that is found to be nonclassical is broader than the set that is found when one does not assume linearity. For instance, there are processes on unipartite systems that are judged to be nonclassical.

One may also wonder how our results are consistent with certain foils to quantum theory, such as the stabilizer subtheory of qutrits and Gaussian quantum mechanics, which exhibit entanglement, incompatibility, and steering despite being classically explainable [6,112] (i.e., despite being consistent with a generalized-noncontextual ontological model). This does not conflict with our results because these foil theories are subtheories of quantum theory that contain a strict subset of the states and measurements allowed in quantum theory, and the notion of classicality proposed here depends on considering all possible dual processes allowed in the theory, which will be different in these subtheories compared to quantum theory.

Extending our analysis of nonclassicality from quantum theory to more general GPTs, and using this to characterize different theories—in the spirit of the GPT reconstructions of quantum theory—is a natural and interesting direction for future work.

Several previous works [33,113,114] also studied the properties of a bipartite state that are pertinent to its ability to support a proof of generalized contextuality, in the following sense. In a steering scenario for a bipartite system A_1A_2 , if one considers implementing any local measurement on A_1 , the bipartite state on A_1A_2 induces a multi-source on A_2 . This multi-source, together with an arbitrary measurement on A_2 constitutes a prepare-measure circuit, and one can ask whether the statistics of this prepare-measure circuit admit of a noncontextual ontological model. In particular, these prior works have studied the connection between these sorts of properties of a bipartite state and their entanglement properties, steering properties, and nonlocality properties.

However, none of these works considered the full set of operational identities in the bipartite scenario [listed in Eq. (42)], which is essential if one wishes to define or fully characterize the classicality of a bipartite state in the sense of finding necessary and sufficient conditions. Essentially, all these works consider the prepare-measure scenario induced by steering from one party to another. Consequently, they consider only the operational identities $\mathcal{O}(M_1 \circ P)$ and $\mathcal{O}(M_2)$ (or a subset thereof), while omitting the identities that arise when considering the preparation steered in the opposite direction, from A_2 to A_1 , namely, $\mathcal{O}(M_2 \circ P)$ and $\mathcal{O}(M_1)$, and while also omitting the identities $\mathcal{O}(\tilde{P})$ defined in Eq. (38).

To be more specific, the work by Tavakoli and Uola [115] considers only the operational identities in $\mathcal{O}(M_2)$ and a subset of the operational identities in $\mathcal{O}(M_1 \circ P)$, namely, those implied by the no-signaling principle. The results of Wright and Farkas [57] rely on a still smaller set of operational identities, including those following from no signaling, together with a subset of the operational identities in $\mathcal{O}(M_2)$, specifically those arising from the normalization condition $\sum_b M_{b|y} = \sum_b M_{b|y'}$. Plávala and Gühne [113], in comparison, do consider all operational

identities in the prepare-measure scenario induced by steering from A_1 to A_2 , but they do not consider all operational identities on the full bipartite scenario, in particular, not the operational identities $\mathcal{O}(\tilde{P})$ defined in Eq. (38).

Because these earlier works impose fewer constraints on the ontological model than we do here, the results they derive do not fully characterize the classical-nonclassical divide for bipartite states according to our proposal. This is consistent with statements in many of those works, which explicitly state that the results merely provide necessary but not sufficient conditions for noncontextual realizability. The one exception to this is Theorem 2 in Ref. [33], whose statement suggests that it leverages all ontological constraints that follow from the assumption of noncontextuality in the scenario under study, when in fact it does not. We discuss this further in Ref. [72].

In a companion paper [73], we also consider the sufficient conditions for nonclassicality implied by different subsets of the set of all constraints that arise from noncontextuality in bipartite Bell scenarios, and we show that for some particularly simple choices of these subsets, one singles out local, unsteerable, and entangled states.

X. CONCLUSION

In this work, we proposed a definition of the classical-nonclassical divide for individual quantum processes. There are many avenues for future work.

If one seeks only sufficient conditions for nonclassicality of a given process, then rather than considering the classical explainability of the set of circuits induced by composing it with all dual processes, it suffices to show the failure of classical explainability for a subset of dual processes. One can consider the partial order of such subsets as a means to understand what is required to witness nonclassicality. For instance, one could seek to find the least elements in this partial order.

Another important question which we study in forthcoming work [116] is how classicality behaves under composition. For example, we prove therein that the composition of two classical processes (either in parallel or in sequence) is always classical; that is, there exists no superactivation of nonclassicality.

This work focused on the decision problem regarding the classical-nonclassical divide: assessing whether a given process is classical or nonclassical. An important open problem is that of quantifying the degree of nonclassicality according to the approach introduced here.

In a companion paper [72], we give some basic tools for doing so for multi-sources and multi-measurements, based on specific measures of how robust a given process is to noise.

The most principled and informative approach to the quantification problem, however, is to pursue it within the framework of resource theories [117]. Roughly speaking, one expects that the correct choice of the divide between

free and nonfree resources in the resource theory should follow from the divide between classical and nonclassical processes.

One can already recognize some aspects that such a resource theory will satisfy. For example, it will not be convex, since (for example) any set of product states is classical, while some separable states are nonclassical. Similarly, the convex combination of two classical measurements can result in a nonclassical measurement. One way to understand this nonconvexity is to note that the process of mixing can introduce new operational identities, and hence new constraints on the ontological model, which in turn can activate nonclassicality.

Moreover, this work opens new possibilities for constructing a resource theory of nonclassicality based on nonrealizability by a noncontextual ontological model; for instance, it could be used to define a resource theory in which the objects under study (the resources) are themselves individual quantum processes. Prior to this work, it seemed that the resources would need to be more complicated mathematical objects (such as sets of states together with sets of effects, or sets of states together with operational equivalences among these; see, for example, Ref. [67]).

An example of a potential application of such a resource theory is to give a resource-theoretic account of the simultaneous self-testing of the nonclassicality in the bipartite state and in the local measurements of a Bell scenario. This cannot be achieved in the LOSR approach to quantifying nonclassicality, because although it can yield a resource-theoretic account of self-testing for the bipartite state, as described in Ref. [4], it cannot yield a resource-theoretic characterization of self-testing for the local measurements, since all local measurements are taken to be free in the LOSR resource theory.

In the Introduction, we noted that the problem of understanding the scope of quantum-over-classical advantages for various tasks motivates understanding the classical-nonclassical divide. In recent years, it has become apparent that the framework of resource theories provides a particularly powerful perspective on this problem. Understanding quantum-over-classical advantages, therefore, provides another motivation for developing a unified resource theory of nonclassicality, i.e., one wherein the notion of freeness is defined in a manner that can be particularized to any type of process. Furthermore, by leveraging the similarity of the structure of nonclassical resources for processes of different types (for instance, multi-sources and multi-measurements), a known quantum-over-classical advantage in one task might be used to infer an advantage in a distinct task.

ACKNOWLEDGMENTS

We thank two anonymous referees for catching some mistakes in an earlier version of the article and for

comments that led us to make many improvements in the presentation. R. W. S. thanks Gilad Gour and Teiko Heinosaari for early discussions on whether incompatibility of measurements should be viewed as a type of nonclassicality. We thank John Selby, Ana Belén Sainz, Roberto D. Baldijão, Rafael Wagner, and Martin Plávala for helpful discussions. Y. Z. acknowledges support from the Institute for Quantum Computing, which is supported by Innovation, Science and Economic Development Canada. Y. Z. is grateful for the hospitality of Perimeter Institute, where this work was carried out. D. S., R. W. S., and Y. Y. were supported by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported in part by the Government of Canada through the Department of Innovation, Science and Economic Development and by the Province of Ontario through the Ministry of Colleges and Universities. Y. Y. was also supported by the Natural Sciences and Engineering Research Council of Canada (Grant No. RGPIN-2024-04419).

DATA AVAILABILITY

No data were created or analyzed in this study.

APPENDIX A: FRAME REPRESENTATIONS

Consider a complex Hilbert space $\mathcal{H}$ with some finite dimension d . The Hermitian operators $\text{Herm}(\mathcal{H})$ on $\mathcal{H}$ form a real vector space of dimension d^2 with the Hilbert-Schmidt inner product $\langle A, B \rangle := \text{Tr}[AB]$ for $A, B \in \text{Herm}(\mathcal{H})$.

Therefore, some standard vector space concepts can be applied to the operator space. For instance, it is well known that a basis of a vector space uniquely defines a dual basis in the dual vector space (i.e., the space of functionals on the vector space). We will leverage this fact, but in a slightly different form.

Lemma 4. Given any basis $\{F_\lambda\}_\lambda$ for a d^2 -dimensional real vector space $\text{Herm}(\mathcal{H})$ of Hermitian operators on a Hilbert space $\mathcal{H}$, there is a unique set $\{D_\lambda\}_\lambda$ of d^2 Hermitian operators satisfying

$$\text{Tr}(D_{\lambda'} F_\lambda) = \delta_{\lambda, \lambda'}, \quad (\text{A1})$$

and $\{D_\lambda\}_\lambda$ also forms a basis for $\text{Herm}(\mathcal{H})$.

Proof. Consider any basis $\{F_\lambda\}_\lambda$ of $\text{Herm}(\mathcal{H})$. It uniquely specifies a basis $\{\mathcal{D}_\lambda\}_\lambda$ of the dual vector space $\text{Herm}(\mathcal{H})^*$, where $\{\mathcal{D}_\lambda\}_\lambda$ are linear functionals satisfying $\mathcal{D}_{\lambda'}(F_\lambda) = \delta_{\lambda, \lambda'}$. To see that this is unique, consider a linear functional $\mathcal{D}'_{\lambda'}$ satisfying $\mathcal{D}'_{\lambda'}(F_\lambda) = \delta_{\lambda, \lambda'}$ for all λ . Since a linear functional is fully specified by its action on a basis, $\mathcal{D}'_{\lambda'}$ is the exact same functional as $\mathcal{D}_{\lambda'}$. Now, in order to obtain again a set of Hermitian operators $\{D_\lambda\}_\lambda$, we use the Riesz representation Theorem [118], which states that each of these linear functionals $\mathcal{D}_\lambda$ can be written as the Hilbert-Schmidt inner product with a unique Hermitian operator

D_λ , namely, $\mathcal{D}_\lambda(\cdot) = \text{Tr}[(\cdot)D_\lambda]$. This picks out a unique basis $\{D_\lambda\}_\lambda$ which satisfies Eq. (A1). ■

More generally, one can define a frame and a dual frame as a generalization of the notion of a basis and a dual basis [119], where the vectors in both the frame and the dual frame are assumed to be spanning, but may or may not be linearly independent [119]. If $\{F_\lambda\}_\lambda$ and $\{D_\lambda\}_\lambda$ denote a frame and a dual frame in $\text{Herm}(\mathcal{H})$, respectively, then it follows that, for any $A \in \text{Herm}(\mathcal{H})$, one has

$$\begin{aligned} A &= \sum_\lambda \text{Tr}[AD_\lambda]F_\lambda, \\ A &= \sum_\lambda \text{Tr}[AF_\lambda]D_\lambda. \end{aligned} \quad (\text{A2})$$

These reconstruction equations show explicitly how every vector in the space can be expanded in terms of the spanning frame operators or the spanning dual-frame operators. Note that, unlike dual bases, dual frames are not necessarily unique. A detailed exploration of frames, dual frames, and their applications in quantum information science lies beyond the scope of this paper. For a more comprehensive introduction, we refer the reader to Ref. [119].

APPENDIX B: OPERATIONAL IDENTITIES IN A QUANTUM CIRCUIT

In Sec. II C of the main text, we discussed the classical explainability of a given prepare-transform-measure circuit, an analysis that relies on enumerating all operational identities in the quantum circuit. In this appendix, we make explicit the assumptions underlying that enumeration and justify why, under these assumptions, the operational identities listed in Sec. II C are the only ones that are relevant. We also briefly explain how this family of operational identities would change under alternative modeling choices, and why we do not adopt those alternatives in our analysis.

When assessing classical explainability, one must account for all operational identities implied by linearity and diagram preservation. Concretely, one should (i) consider all quantum processes obtainable via finite parallel and sequential compositions of the set of processes under consideration, since, in principle, each such composite process can be inferred from the experimental statistics, and (ii) group these processes according to their quantum input and output system types, since an operational equivalence condition can hold only among processes that have the same quantum inputs and outputs.

The prepare-transform-measure scenario discussed in Sec. II C contains a multi-state on system A : $\{\rho_{|x}^A\}_x$, a multi-channel from system A to B : $\{\mathcal{E}_{|z}^{B|A}\}_z$, and a measurement on system B : $\{M_b^B\}_b$, as depicted in Fig. 3. Considering arbitrary compositions of these processes

leads one to consider also a multi-state on system B : $\{\tilde{\rho}_{|xz}^B\}_{xz}$, a multi-measurement on system A : $\{\{\tilde{M}_{b|z}^A\}_b\}_z$, and a measure-and-prepare instrument from system A to B : $\{\{\tilde{\mathcal{E}}_{b''|z'x'z''}^{B|A}\}_b\}_{xz}$. The operational identities among these processes are described in Eqs. (3), (5), (7), (10), and (12) in Sec. II C.

Under the assumption that A and B are distinct system types, the list above is indeed the set of all operational identities. In particular, it does not include any that refer to measure-and-prepare channels constructed by composing states $\rho_{|x}^A$ and effects M_b^B . This is the case since the only operational identities on these measure-and-prepare channels from B to A are implied by those holding among the individual states and among the individual effects. Unlike in the case of channels from A to B , there are no other possible channels from B to A within this scenario that could be combined with these measure-and-prepare channels to lead to new operational identities.

In principle, one could also consider the case where A and B have the same system type. In this case, additional operational identities can arise, for example, relating ρ_x^A and $\mathcal{E}_{|z}^{B|A}(\rho_{|x}^A)$, M_b^A and $[\mathcal{E}_{|z}^{B|A}]^\dagger(M_b^A)$, or involving sequential compositions of the channels themselves (note that such sequential compositions are not possible if the system types A and B are distinct).

Throughout the main text, we assume that any two distinct quantum inputs and outputs belong to different system types. This choice is motivated by the problem of gauge freedom from GPT tomography and gate-set tomography [73,120,121]. Specifically, in self-consistent tomography, one does not obtain a unique output of the processes in a quantum circuit; rather, one obtains an equivalence class of representations related by invertible linear transformations on each system, which are referred to as tomographic gauge freedoms, following Ref. [121]. Without additional assumptions beyond the circuit data, there is no canonical identification between the representations for different quantum systems; therefore, treating two systems as the same “type” is itself a gauge choice.

Thus, by treating distinct quantum systems in the circuit as distinct types, we restrict ourselves to operational identities that are gauge independent. Therefore, our result can be seen as a gauge-independent assessment of the nature of the process, which is readily experimentally implementable using schemes similar to various self-consistent tomographies [73,120,121].

APPENDIX C: WHY THE FORM OF VARIABILITY OVER DUAL PROCESSES IS IRRELEVANT

In the main text (especially Sec. II D), we noted that there are many ways of ranging over all quantum states. For example, one can range over all quantum states using a

multi-state process, where the setting variable ranges over all possible quantum states, or by postselecting on the outcome of a source associated with the unitarily covariant ensemble of states, or via a multi-source with both setting and outcome variables. We noted that the different choices one can make in this regard make no difference for qualitative assessments of nonclassicality. This is because every multi-source defines a collection of (possibly subnormalized) quantum states, and the ontological representation of a collection of states defines the associated collection of representations of the states; moreover, by the linearity of any ontological model, one can uniquely extend the ontological representation of the normalized states to an ontological representation of subnormalized versions of those same states.

The story is similar for measurements. An ontological model for some set of measurements constitutes a collection of representations of all of the effects appearing in those measurements. These response functions serve to define a valid and empirically adequate ontological model for any other set of measurements that contains the same set of effects. Consequently, the existence of a linear and empirically adequate mapping from the set of all quantum effects to their ontological representation as response functions implies the existence of such a mapping for any multi-measurement one could construct with those effects.

It is worth noting that this observation also explains why verdicts of nonclassicality are unchanged under flag convexification (see Lemma 2). Whether one fixes a certain input setting value or postselects on the corresponding flag variable taking that same value, one is simply choosing between two different physical instantiations of the same variability, so an ontological representation for one directly gives an ontological representation for the other.

APPENDIX D: PROOF OF THE SUFFICIENCY OF CONSIDERING FACTORIZING DUALS

We now prove Theorem 1, that the definition of nonclassicality obtained by considering all dual processes and the one obtained by considering all factorizing dual processes is the same. In other words, we prove that Definition 3 is equivalent to Definition 3'.

In the following two subsections, we prove the result for bipartite states and for transformations, respectively. The generalization of the proof to arbitrary process types is straightforward.

Although the proof is lengthy, its basic idea is simple. First, we note that every quantum process with multipartite inputs and/or outputs can be written as a linear combination of processes that factorize across the inputs and outputs. From this and the assumption that an ontological model is diagram preserving and linear, one can then uniquely extend any given ontological model defined for all product processes in quantum theory to a quasiprobabilistic representation of all processes in quantum theory,

including the nonproduct ones. It remains only to show that for the relevant processes in question, this quasiprobabilistic representation is in fact a valid probabilistic representation (e.g., is positive and normalized appropriately), a fact which follows easily from frame-representation theory [122].

1. Factorizing duals for bipartite states

We first prove the theorem for the special case where the given process is a bipartite state. That is, we prove that the definition of nonclassicality for a bipartite state is the same whether it is defined relative to all dual processes or to all factorizing dual processes.

Lemma 5. The statistics generated by a bipartite state together with all bipartite effects are classically explainable (in the sense of Definition 1) if and only if the statistics generated by it with all product effects are classically explainable.

Proof. The proof for the only if direction is simply that factorizing product bipartite effects are special cases of general bipartite effects. So if the statistics generated by a bipartite state with all bipartite effects are classically explainable, then those generated by that state with all product effects are classically explainable.

Now we prove the if direction. If the statistics generated by a bipartite state $\rho^{A_1 A_2}$ with every product bipartite effect $M^{A_1} \otimes M^{A_2}$ are classically explainable, then there exists an ontological model for the statistics such that

$$\text{Tr}[M^{A_1} \otimes M^{A_2} \rho^{A_1 A_2}] = \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) p(M^{A_1} | \lambda_1) p(M^{A_2} | \lambda_2), \quad (\text{D1})$$

where $p(M^{A_i} | \lambda)$ depends linearly on the effect $M^{A_i} \in \mathcal{M}(\mathcal{H}^{A_i})$. Moreover since $p(M^{A_i} | \lambda) \in [0, 1]$ and $p(\mathbb{1}^{A_i} | \lambda) = 1$, by the generalized Gleason theorem [123], there exists a unique set of density operators $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i}$ such that $p(M^{A_i} | \lambda_i) = \text{Tr}[M^{A_i} \sigma_{\lambda_i}^{A_i}]$, and so we have

$$\begin{aligned} \text{Tr}[M^{A_1} \otimes M^{A_2} \rho^{A_1 A_2}] &= \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) \text{Tr}[M^{A_1} \sigma_{\lambda_1}^{A_1}] \text{Tr}[M^{A_2} \sigma_{\lambda_2}^{A_2}] \\ &= \text{Tr}[(M^{A_1} \otimes M^{A_2}) \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}] \end{aligned} \quad (\text{D2})$$

for any $M^{A_i} \in \mathcal{M}(\mathcal{H}^{A_i})$. Because the set of all product effects is informationally complete, Eq. (D2) implies that

$$\rho^{A_1 A_2} = \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}, \quad (\text{D3})$$

and so the state $\rho^{A_1 A_2}$ must be separable.

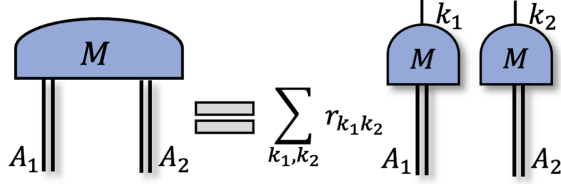


FIG. 25. In quantum theory, every effect can be decomposed as a linear combination of product effects.

In quantum theory, any bipartite effect $M^{A_1 A_2} \in \mathcal{M}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{A_2})$ can always be decomposed into a linear combination of product effects (see Fig. 25) as

$$M^{A_1 A_2} = \sum_{k_1, k_2} r_{k_1 k_2} M_{k_1}^{A_1} \otimes M_{k_2}^{A_2} \quad (\text{D4})$$

with some real coefficients $\{r_{k_1 k_2}\}_{k_1 k_2}$. Consequently, we can define a unique quasiprobability representation for $M^{A_1 A_2}$ as the unique linear extension of the ontological representation of the product effects via

$$\begin{aligned} p(M^{A_1 A_2} | \lambda_1 \lambda_2) &:= \sum_{k_1 k_2} r_{k_1 k_2} p(M_{k_1}^{A_1} | \lambda_1) p(M_{k_2}^{A_2} | \lambda_2) \\ &= \sum_{k_1, k_2} r_{k_1 k_2} \text{Tr}[M_{k_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \text{Tr}[M_{k_2}^{A_2} \sigma_{\lambda_2}^{A_2}] \\ &= \text{Tr}[M^{A_1 A_2} \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}]. \end{aligned} \quad (\text{D5})$$

The form of Eq. (D5) immediately guarantees that the representation map is linear and diagram preserving. Furthermore, since $\sigma_{\lambda_1}^{A_1}$ and $\sigma_{\lambda_2}^{A_2}$ are valid quantum states, $p(M^{A_1 A_2} | \lambda_1 \lambda_2) \in [0, 1]$ defines a valid stochastic representation.

Finally, it is straightforward to verify that the ontological model for the given effect and state $\rho^{A_1 A_2}$ reproduces the quantum statistics

$$\begin{aligned} \sum_{\lambda_1 \lambda_2} p(M^{A_1 A_2} | \lambda_1 \lambda_2) p(\lambda_1 \lambda_2) &= \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) \sum_{k_1 k_2} r_{k_1 k_2} \text{Tr}[M_{k_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \text{Tr}[M_{k_2}^{A_2} \sigma_{\lambda_2}^{A_2}] \\ &= \text{Tr}[M^{A_1 A_2} \rho^{A_1 A_2}]. \end{aligned} \quad (\text{D6})$$

Consequently, any statistics generated by $\rho^{A_1 A_2}$ with any bipartite effect (not necessarily factorizing) are classically explainable. This concludes the proof of the if direction. ■

2. Factorizing duals for transformations

Next, we prove that the statistics for the set of circuits generated by contracting the transformation of interest with

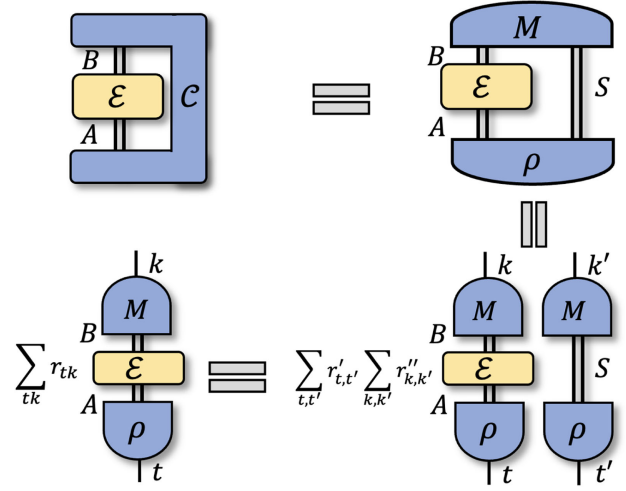


FIG. 26. A comb can be decomposed using state-effect pairs.

all dual combs are classically explainable if and only if this is the case with all factorizing dual combs.

Before proving this, we include an aside about the ontological representation of quantum combs. To our knowledge, this has not been explicitly discussed in the literature. Nevertheless, the fully compositional definition of the ontological representation map given in Ref. [9] can be applied to this special case, and we do so here.

We consider the specific class of combs which are dual processes for a given transformation $\mathcal{E}_{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$, as shown in the top left of Fig. 26. These combs are denoted $\mathcal{C}^{A|B}$. $\mathcal{C}^{A|B}$ is a special case of 2-combs defined in Ref. [83], which has only a quantum output on $\mathcal{L}(\mathcal{H}^A)$ and a quantum input on $\mathcal{L}(\mathcal{H}^B)$.

An ontological model of the circuit in Fig. 26 associates with $\mathcal{E}$ a stochastic map $p_{\mathcal{E}}(\lambda_B | \lambda_A)$ and with $\mathcal{C}$ a substochastic map $\tilde{p}_{\mathcal{C}}(\lambda_A | \lambda_B)$ [note that in our case, a comb is defined by a single effect so that its ontological model is subnormalized, i.e., $\sum_{\lambda_A} \tilde{p}_{\mathcal{C}}(\lambda_A | \lambda_B)$ need not to be 1]. As always for an ontological model, the map from the set of quantum processes of a given type (in this case, from the set of combs) to their ontological representations must be linear, and moreover, the composition of the processes within quantum theory must equal the composition of their ontological representations (so that the empirical predictions of quantum theory are reproduced). Linearity implies that the $\tilde{p}_{\mathcal{C}}(\lambda_A | \lambda_B)$ obey the constraints implied by all operational identities among combs, while the latter fact implies that

$$\mathcal{C}^{A|B} \circ \mathcal{E}^{B|A} = \sum_{\lambda_A \lambda_B} \tilde{p}_{\mathcal{C}}(\lambda_A | \lambda_B) p_{\mathcal{E}}(\lambda_B | \lambda_A). \quad (\text{D7})$$

We can now give the main argument. Note that a factorizing comb with quantum output A and input B is simply a state on A and an effect on B .

Lemma 6. The statistics of the set of circuits obtained by composing a quantum channel together with every possible comb is classically explainable if and only if the statistics of the set of circuits obtained by composing it with every possible factorizing comb is classically explainable.

Proof. The only if direction is trivial, as a factorizing comb is a special case of a comb.

Now we prove the if direction. If the statistics generated by the circuit consists of a quantum channel $\mathcal{E}^{B|A}$ and a factorizing comb defined by a state ρ^A and effect M^B that are classically explainable, then we have an ontological model for the operational statistics

$$\text{Tr}[M^B \mathcal{E}^{B|A}(\rho^A)] = \sum_{\lambda_A \lambda_B} p(M^B | \lambda_B) p(\lambda_B | \lambda_A) p(\lambda_A | \rho^A), \quad (\text{D8})$$

where $p(M^B | \lambda_B)$ depends linearly on M^B and $p(\lambda_A | \rho^A)$ depends linearly on all ρ^A , respectively. Since $p(M^B | \lambda_B) \in [0, 1]$ with $p(\mathbb{1}^B | \lambda_B) = 1$, from the generalized Gleason theorem (and the Riesz representation theorem), there exists a unique POVM $\{G_{\lambda_A}^A\}_{\lambda_A}$ such that $p(\lambda_A | \rho^A) = \text{Tr}[\rho^A G_{\lambda_A}^A]$ and that there exists a unique set of density operator $\{\sigma_{\lambda_B}^B\}_{\lambda_B}$ such that $p(M^B | \lambda_B) = \text{Tr}[M^B \sigma_{\lambda_B}^B]$. Therefore,

$$\begin{aligned} \text{Tr}[M^B \mathcal{E}^{B|A}(\rho^A)] &= \sum_{\lambda_A \lambda_B} p(\lambda_B | \lambda_A) \text{Tr}[\rho^A G_{\lambda_A}^A] \text{Tr}[M^B \sigma_{\lambda_B}^B] \\ &= \text{Tr} \left[M^B \sum_{\lambda_A \lambda_B} p(\lambda_B | \lambda_A) \text{Tr}[G_{\lambda_A}^A \rho^A] \sigma_{\lambda_B}^B \right]. \end{aligned} \quad (\text{D9})$$

Since this holds for all quantum effects $M^B \in \mathcal{M}(\mathcal{H}^B)$, we have

$$\mathcal{E}^{B|A}(\rho^A) = \sum_{\lambda_A \lambda_B} p(\lambda_B | \lambda_A) \text{Tr}[G_{\lambda_A}^A \rho^A] \sigma_{\lambda_B}^B \quad (\text{D10})$$

for all $\rho^A \in \mathcal{S}(\mathcal{H}^A)$, which means that $\mathcal{E}^{B|A}$ can be expressed in the Holevo form and therefore must be entanglement breaking [124].

Now, we will use these ontological representations to show that we can also construct an ontological representation of the comb $\mathcal{C}$.

First, we will rewrite the comb $\mathcal{C}^{A|B}$ within quantum theory as a linear combination of factorizing combs as shown in Fig. 26; then, we will use the above representations of the states and effects appearing in the factorizing comb to construct a quasiprobabilistic representation of $\mathcal{C}^{A|B}$. This representation is then proven to be non-negative and therefore to define a valid ontological representation. The detailed steps are as follows.

Specifically, it has been shown [83] that any comb $\mathcal{C}^{A|B}$ can be decomposed as a bipartite state ρ^{AS} followed by a

bipartite measurement (effect) M^{BS} with an identity channel acting on system S . That is,

$$\mathcal{C}^{A|B}(\mathcal{E}^{B|A}) := \text{Tr}[M^{BS}(\mathcal{E}^{B|A} \otimes \mathbb{1}^{S|S})\rho^{AS}]. \quad (\text{D11})$$

Furthermore, any bipartite quantum state and bipartite measurement can be decomposed as

$$\rho^{AS} = \sum_{tt'} r'_{tt'} \rho_t^A \otimes \rho_{t'}^S, \quad (\text{D12a})$$

$$M^{BS} = \sum_{kk'} r''_{kk'} M_k^B \otimes M_{k'}^S \quad (\text{D12b})$$

with some real coefficients $\{r'_{tt'}\}_{tt'}$ and $\{r''_{kk'}\}_{kk'}$. As such, the action of the comb $\mathcal{C}^{A|B}$ on the channel $\mathcal{E}^{B|A}$ can be further decomposed as

$$\begin{aligned} \mathcal{C}^{A|B}(\mathcal{E}^{B|A}) &:= \text{Tr} \left[\left(\sum_{kk'} r''_{kk'} M_k^B \otimes M_{k'}^S \right) (\mathcal{E}^{B|A} \otimes \mathbb{1}) \right. \\ &\quad \left. \times \left(\sum_{tt'} r'_{tt'} \rho_t^A \otimes \rho_{t'}^S \right) \right]. \end{aligned} \quad (\text{D13})$$

Defining the new real coefficients

$$r_{tk} := \sum_{tt'kk'} r'_{tt'} r''_{kk'} \text{Tr}[M_{k'}^S \rho_{t'}^S], \quad (\text{D14})$$

the expression for the comb $\mathcal{C}^{A|B}$ acting on $\mathcal{E}^{B|A}$ can be further simplified to

$$\mathcal{C}^{A|B}(\mathcal{E}^{B|A}) = \sum_{tk} r_{tk} \text{Tr}[M_k^B \mathcal{E}^{B|A}(\rho_t^A)]. \quad (\text{D15})$$

This final decomposition of the comb $\mathcal{C}^{A|B}$ is shown on the bottom left diagram of Fig. 26.

Now, given our choice to represent states and effects as in Eq. (D9), the unique linear and diagram-preserving representation of the comb $\mathcal{C}^{A|B}$ can be defined as

$$\begin{aligned} \tilde{p}_C(\lambda_A | \lambda_B) &:= \sum_{tk} r_{tk} p(\lambda_A | \rho_t^A) p(M_k^B | \lambda_B) \\ &= \sum_{tk} r_{tk} \text{Tr}[G_{\lambda_A}^A \rho_t^A] \text{Tr}[M_k^B \sigma_{\lambda_B}^B] \\ &= \text{Tr}[G_{\lambda_A}^A \mathcal{C}^{A|B}(\sigma_{\lambda_B}^B)]. \end{aligned} \quad (\text{D16})$$

Furthermore, since $G_{\lambda_A}^A$ are valid effect and $\sigma_{\lambda_B}^B$ are valid quantum states, $\tilde{p}_C(\lambda_A | \lambda_B) \in [0, 1]$ defines a valid substochastic representation.

Finally, it is straightforward to verify that these representations reproduce the statistics

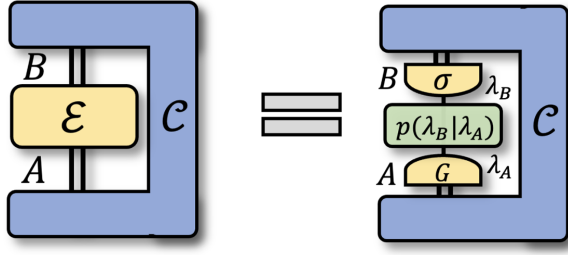


FIG. 27. An entanglement-breaking channel and a dual.

$$\sum_{\lambda_A \lambda_B} \tilde{p}_C(\lambda_A | \lambda_B) p_E(\lambda_B | \lambda_A) = \sum_{\lambda_A \lambda_B} p_E(\lambda_B | \lambda_A) \text{Tr}[G_{\lambda_A}^A C^{A|B}(\sigma_{\lambda_B}^B)] = \mathcal{E}^{B|A} \circ \mathcal{C}^{A|B}, \quad (\text{D17})$$

where in the second equality, we used the Holevo form of the channel $\mathcal{E}_{B|A}$ from Eq. (D10) (also see Fig. 27).

Thus, $\tilde{p}_C(\lambda_A | \lambda_B)$ is a valid ontological representation for a generic quantum comb $\mathcal{C}^{A|B}$. That is, we have seen that whenever the statistics generated by $\mathcal{E}^{B|A}$ with all factorizing pairs of effect and state are classically explainable, one can write down such a representation for any comb $\mathcal{C}^{A|B}$ (possibly nonfactorizing). It follows that the statistics generated by $\mathcal{E}^{B|A}$ with any comb must also be classically explainable. This concludes the if direction of the proof. ■

3. Factorizing duals for arbitrary process

With the two special cases discussed in the previous subsections, we now prove our main theorem analogously for the most general quantum process: a multipartite multi-instrument.

Theorem 1. A quantum process is such that there exists some set of dual processes that it can be contracted with to obtain statistics that are not classically explainable, if and only if there exists some set of factorizing dual processes for which this is the case.

Proof. The core idea of the proof is similar to the ones for the if direction in Lemmas 5 and 6. That is, if one can construct an ontological model for the given quantum process contracted with every possible factorizing dual process, then one can obtain an ontological representation for every dual process (possibly nonfactorizing) by building it from the representations of the factorizing dual processes in a linear, diagram-preserving manner.

In this paper, the most general process we consider is a multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\bar{B}|\bar{A}}\}_c\}_z: \mathcal{L}(\otimes_{i=1}^n \mathcal{H}^{A_i}) \mapsto \mathcal{L}(\otimes_{j=1}^m \mathcal{H}^{B_j})$; i.e., each $\mathcal{F}_{c|z}^{\bar{B}|\bar{A}}$ is a CP trace-nonincreasing map such that $\sum_c \mathcal{F}_{c|z}^{\bar{B}|\bar{A}}$ defines the quantum channel.

If the statistics generated by a multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\bar{B}|\bar{A}}\}_c\}_z$ is classical for factorizing duals

composed of the set of states $\{\rho^{A_i}\}$ and the set of effects $\{M^{B_j}\}$, then there necessarily exists an ontological model for the operational statistics such that

$$\begin{aligned} \text{Tr} \left[\left(\bigotimes_{j=1}^m M^{B_j} \right) \mathcal{F}_{c|z}^{\bar{B}|\bar{A}} \left(\bigotimes_{i=1}^n \rho^{A_i} \right) \right] \\ = \sum_{\lambda_{\bar{B}} \lambda_{\bar{A}}} p(c \lambda_{\bar{B}} | z \lambda_{\bar{A}}) \prod_j p(M^{B_j} | \lambda_{B_j}) \prod_i p(\lambda_{A_i} | \rho^{A_i}) \\ = \sum_{\lambda_{\bar{B}} \lambda_{\bar{A}}} p(c \lambda_{\bar{B}} | z \lambda_{\bar{A}}) \prod_j \text{Tr}[M^{B_j} \sigma_{\lambda_{B_j}}^{B_j}] \prod_i \text{Tr}[G_{\lambda_{A_i}}^{A_i} \rho^{A_i}], \quad (\text{D18}) \end{aligned}$$

where in the second equality we write $p(M^{B_j} | \lambda_{B_j}) = \text{Tr}[M^{B_j} \sigma_{\lambda_{B_j}}^{B_j}]$ and $p(\lambda_{A_i} | \rho^{A_i}) = \text{Tr}[G_{\lambda_{A_i}}^{A_i} \rho^{A_i}]$ using Gleason's theorem and the Riesz representation theorem since these ontological processes depend linearly on all quantum effects M^{B_j} and all quantum states ρ^{A_i} , respectively. Because the set of all product effects and states is informationally complete for characterizing the multipartite multi-instrument, it follows that

$$\mathcal{F}_{c|z}^{\bar{B}|\bar{A}}(\cdot) = \sum_{\lambda_{\bar{B}} \lambda_{\bar{A}}} p(c \lambda_{\bar{B}} | z \lambda_{\bar{A}}) \bigotimes_j \sigma_{\lambda_{B_j}}^{B_j} \text{Tr} \left[\bigotimes_i G_{\lambda_{A_i}}^{A_i}(\cdot) \right]. \quad (\text{D19})$$

Similar to the construction of ontological models for general dual processes in Lemmas 5 and 6, we can now proceed to define the ontological model for a general dual process $\mathcal{C}^{\bar{A}|\bar{B}}: \mathcal{L}(\otimes_{i=1}^n \mathcal{H}^{A_i}) \mapsto \mathcal{L}(\otimes_{j=1}^m \mathcal{H}^{B_j})$ for the target process $\{\{\mathcal{F}_{c|z}^{\bar{B}|\bar{A}}\}_c\}_z$. To do so, we will need to first find a decomposition of the general dual process $\mathcal{C}^{\bar{A}|\bar{B}}$ (a multipartite comb) in terms of product states and product effects as we did in Figs. 25 and 26. Such a decomposition always exists because quantum theory satisfies the principle of tomographic locality. That is, for any channel $\mathcal{E}^{\bar{B}|\bar{A}}$, we have

$$\mathcal{C}^{\bar{A}|\bar{B}} \circ \mathcal{E}^{\bar{B}|\bar{A}} := \sum_{\bar{r} \bar{k}} r_{\bar{r} \bar{k}} \text{Tr} \left[\bigotimes_j M_{k_j}^{B_j} \mathcal{E}^{\bar{B}|\bar{A}} \left(\bigotimes_i \rho_{t_i}^{A_i} \right) \right] \quad (\text{D20})$$

for some real coefficients $\{r_{\bar{r} \bar{k}}\}$. Now, given the ontological representation we have for states and effects in Eq. (D18), we can construct the ontological representations for the general dual process $\mathcal{C}^{\bar{A}|\bar{B}}$ as

$$\begin{aligned} \tilde{p}_C(\lambda_{\bar{A}} | \lambda_{\bar{B}}) &:= \sum_{\bar{r} \bar{k}} r_{\bar{r} \bar{k}} \prod_j p(M_{k_j}^{B_j} | \lambda_{B_j}) \prod_i p(\lambda_{A_i} | \rho_{t_i}^{A_i}) \\ &= \sum_{\bar{r} \bar{k}} r_{\bar{r} \bar{k}} \prod_j \text{Tr}[M_{k_j}^{B_j} \sigma_{\lambda_{B_j}}^{B_j}] \prod_i \text{Tr}[G_{\lambda_{A_i}}^{A_i} \rho_{t_i}^{A_i}] \\ &= \text{Tr} \left[\bigotimes_i G_{\lambda_{A_i}}^{A_i} \mathcal{C}^{\bar{A}|\bar{B}} \left(\bigotimes_i \sigma_{\lambda_{B_j}}^{B_j} \right) \right]. \quad (\text{D21}) \end{aligned}$$

Since $G_{\lambda_{A_i}}^{A_i}$ and $\sigma_{\lambda_{B_j}}^{B_j}$ are valid effects and states, we have $p_C(\lambda_{\vec{A}}|\lambda_{\vec{B}}) \in [0, 1]$ defining a substochastic representation for any valid dual process $\mathcal{C}^{\vec{A}|\vec{B}}$.

Therefore, once one has constructed an ontological representation of a multipartite multi-instrument along with all factorizing combs dual to it, one can uniquely extend this ontological representation to all combs (including those that are nonfactorizing). ■

APPENDIX E: PROOF OF THEOREM 2

We now repeat and then prove Theorem 2.

Theorem 2. For a multi-source $\mathbf{P} := \{\{p(a|x)\rho_{a|x}\}_a\}_x$, the following statements are equivalent:

- (0) It is classical (in the sense of Definition 5).
- (1) Each of the unnormalized states $p(a|x)\rho_{a|x}$ can be decomposed as

$$p(a|x)\rho_{a|x} = \sum_{\lambda} p(a\lambda|x)\sigma_{\lambda}, \quad \forall a, x \quad (\text{E1})$$

for some fixed set of normalized states $\{\sigma_{\lambda}\}_{\lambda}$ and some conditional probability distribution $p(a\lambda|x)$ satisfying

$$\sum_{ax} \alpha_{ax} p(a\lambda|x) = 0, \quad \forall \{\alpha_{ax}\}_{ax} \in \mathcal{O}(\mathbf{P}) \quad (\text{E2})$$

for all λ , where $\mathcal{O}(\mathbf{P})$ is defined in Eq. (17a).

- (2) Each of the normalized states $\rho_{a|x}$ can be decomposed within a frame representation on $V = \text{Span}(\{\rho_{a|x}\}_{ax})$ as

$$\rho_{a|x} = \sum_{\lambda} \text{Tr}[H_{\lambda}\rho_{a|x}]\sigma_{\lambda}, \quad \forall a, x, \quad (\text{E3})$$

where the frame is a fixed set of density operators $\{\sigma_{\lambda}\}_{\lambda}$, and the dual frame is a set of (possibly nonpositive) Hermitian operators, i.e., $\{H_{\lambda}\}_{\lambda} \subset \text{Herm}(\mathcal{H})$ satisfying $\sum_{\lambda} H_{\lambda} = \mathcal{P}_V(\mathbb{1})$ and $\text{Tr}[H_{\lambda}\rho_{a|x}] \in [0, 1]$ for all a, x, λ , with $\mathcal{P}_V$ being the superoperator that projects onto the operator space V . This condition is depicted in Fig. 8.

Proof. We first prove that condition (0) implies (2). From Definition 5, we know that the statistics generated by a classical multi-source $\{\{p(a|x)\rho_{a|x}\}_a\}_x$ with any quantum effect can be expressed as

$$\text{Tr}[M\rho_{a|x}] = \sum_{\lambda} p(M|\lambda)p(\lambda|a, x), \quad \forall a, x, M \in \mathcal{M}(\mathcal{H}),$$

where $\{p(\lambda|a, x)\}_{ax}$ depends linearly on $\{\rho_{a|x}\}_{ax}$ and $p(M|\lambda)$ depends linearly on all quantum effects $M \in \mathcal{M}(\mathcal{H})$.

Since $p(M|\lambda) \in [0, 1]$ with $p(\mathbb{1}|\lambda) = 1$ for unit effect $\mathbb{1}$, we can use the generalized Gleason theorem [123] to infer that $p(M|\lambda) = \text{Tr}[M\sigma_{\lambda}]$ for a unique normalized density matrix σ_{λ} . Similarly, for any linear map from states $\{\rho_{a|x}\}_{ax}$ to probability distributions $\{p(\lambda|a, x)\}_{ax}$, there always exists a Hermitian operator H_{λ} for each λ such that $p(\lambda|a, x) = \text{Tr}[H_{\lambda}\rho_{a|x}]$. Given that $p(\lambda|a, x) \in [0, 1]$ and $\sum_{\lambda} p(\lambda|a, x) = 1$ for all $\rho_{a|x}$, we infer that $\sum_{\lambda} H_{\lambda} = \mathcal{P}_V(\mathbb{1})$ where $\mathcal{P}_V$ is the superoperator that projects $\mathbb{1}$ onto the operator space $V = \text{Span}(\{\rho_{a|x}\}_{a,x})$. We thus have, for any given a and x ,

$$\begin{aligned} \text{Tr}[M\rho_{a|x}] &= \sum_{\lambda} \text{Tr}[M\sigma_{\lambda}]\text{Tr}[H_{\lambda}\rho_{a|x}] \\ &= \text{Tr}\left[M \sum_{\lambda} \text{Tr}[H_{\lambda}\rho_{a|x}]\sigma_{\lambda}\right], \quad \forall M. \end{aligned} \quad (\text{E4})$$

Because Eq. (E5) must hold for all M , it follows that

$$\rho_{a|x} = \sum_{\lambda} \text{Tr}[H_{\lambda}\rho_{a|x}]\sigma_{\lambda}. \quad (\text{E5})$$

To show (2) $\rightarrow$ (1), we can simply define

$$\begin{aligned} p(\lambda|a, x) &:= \text{Tr}[H_{\lambda}\rho_{a|x}], \\ &\Rightarrow p(a\lambda|x) := p(a|x)p(\lambda|a, x) \\ &= \text{Tr}[H_{\lambda}p(a|x)\rho_{a|x}]. \end{aligned} \quad (\text{E6})$$

It follows that $\{p(a\lambda|x)\}_{ax}$ depends linearly on $\{p(a|x)\rho_{a|x}\}_{ax}$ and consequently respects the operational identities $\mathcal{O}(\mathbf{P})$.

To show that condition (1) implies condition (0), we can define $p(M|\lambda) := \text{Tr}[\sigma_{\lambda}M]$, and thus we have

$$\begin{aligned} \text{Tr}[M\rho_{a|x}] &= \sum_{\lambda} \text{Tr}[M\sigma_{\lambda}]p(\lambda|a, x), \quad \forall M, a, x, \\ p(a|x)\text{Tr}[M\rho_{a|x}] &= \sum_{\lambda} \text{Tr}[M\sigma_{\lambda}]p(a\lambda|x), \quad \forall M, a, x, \end{aligned} \quad (\text{E7})$$

where $\{p(a\lambda|x)\}_{ax}$ respects $\mathcal{O}(\mathbf{P})$ and $p(M|\lambda)$ defines a stochastic map that is a linear functional on the set of all quantum effects $M \in \mathcal{M}(\mathcal{H})$. Thus, there exists an ontological model for the statistics generated by the multi-source contracted with any quantum effect; it follows that $\{\{p(a|x)\rho_{a|x}\}_a\}_x$ is classical. ■

APPENDIX F: PROOF OF THEOREM 3

We now repeat and then prove Theorem 3.

Theorem 3. For a multi-measurement $\mathbf{M} = \{\{M_{b|y}\}_b\}_y$, the following statements are equivalent:

- (0) It is classical (in the sense of Definition 6).

(1) Its effects can be decomposed as

$$M_{b|y} = \sum_{\lambda} p(b|y\lambda) G_{\lambda}, \quad \forall b, y \quad (\text{F1})$$

for some POVM $\{G_{\lambda}\}_{\lambda}$ and some conditional probability distribution satisfying

$$\sum_{by} \beta_{by} p(b|y\lambda) = 0, \quad \forall \{\beta_{by}\}_{by} \in \mathcal{O}(\mathbf{M}) \quad (\text{F2})$$

for all λ , where $\mathcal{O}(\mathbf{M})$ is defined in Eq. (17b).

(2) Its effects can be decomposed within a frame representation on $V = \text{Span}(\{M_{b|y}\})$ as

$$M_{b|y} = \sum_{\lambda} \text{Tr}[M_{b|y} F_{\lambda}] G_{\lambda}, \quad \forall b, y, \quad (\text{F3})$$

where $\{G_{\lambda}\}_{\lambda}$ is a POVM with $\sum_{\lambda} G_{\lambda} = \mathcal{P}_V(\mathbb{1}) = \mathbb{1}$ and $\{F_{\lambda}\}_{\lambda} \subset \text{Herm}(\mathcal{H})$ is a set of Hermitian operators satisfying $\text{Tr}[F_{\lambda}] = 1$ and $\text{Tr}[M_{b|y} F_{\lambda}] \in [0, 1]$ for all b, y, λ . This condition is depicted in Fig. 10.

Proof. We first prove that condition (0) implies condition (2). From Definition 6, we know that the statistics generated by a classical multi-measurement $\{\{M_{b|y}\}_b\}_y$ contracted with every possible quantum state can be expressed as

$$\text{Tr}[M_{b|y}\rho] = \sum_{\lambda} p(b|y\lambda) p(\lambda|\rho), \quad \forall b, y, \quad \forall \rho \in \mathcal{S}(\mathcal{H}), \quad (\text{F4})$$

where $\{p(b|y\lambda)\}_{by}$ depends linearly on $\{M_{b|y}\}_{by}$ [and so it respects $\mathcal{O}(\mathbf{M})$], and $p(\lambda|\rho)$ depends linearly on the quantum state $\rho \in \mathcal{S}(\mathcal{H})$.

Using the Riesz representation theorem, we can write this latter linear function as $p(\lambda|\rho) = \text{Tr}[G_{\lambda}\rho]$ for some effect $G_{\lambda} \in \mathcal{M}(\mathcal{H})$. Furthermore, since $\sum_{\lambda} p(\lambda|\rho) = 1$, we must have that $\sum_{\lambda} G_{\lambda} = \mathcal{P}_V(\mathbb{1}) = \mathbb{1}$. Thus, $\{G_{\lambda}\}_{\lambda}$ is a POVM. Similarly, we can always write $p(b|y\lambda) = \text{Tr}[M_{b|y} F_{\lambda}]$ for some Hermitian operator F_{λ} . The fact that $\sum_b p(b|y\lambda) = 1$ implies that $\text{Tr}[F_{\lambda}] = 1$. Thus, we have

$$\begin{aligned} \text{Tr}[M_{b|y}\rho] &= \sum_{\lambda} \text{Tr}[M_{b|y} F_{\lambda}] \text{Tr}[G_{\lambda}\rho] \\ &= \text{Tr} \left[\sum_{\lambda} \text{Tr}[M_{b|y} F_{\lambda}] G_{\lambda} \rho \right], \quad \forall \rho \in \mathcal{S}(\mathcal{H}), \quad (\text{F5}) \end{aligned}$$

where $\text{Tr}[M_{b|y} F_{\lambda}] = p(b|y\lambda) \in [0, 1]$ [as $p(b|y\lambda) \in [0, 1]$] and $\text{Tr}[F_{\lambda}] = 1$ [as $\sum_b p(b|y\lambda) = 1$] for all b, y, λ . Since Eq. (F5) must hold for all ρ , it follows that

$$M_{b|y} = \sum_{\lambda} \text{Tr}[M_{b|y} F_{\lambda}] G_{\lambda}. \quad (\text{F6})$$

To show (2) $\rightarrow$ (1), we can simply define

$$p(b|y\lambda) := \text{Tr}[M_{b|y} F_{\lambda}].$$

It follows that $p(b|y\lambda)$ depends linearly on $M_{b|y}$, which implies that it respects the operational identities in the set $\mathcal{O}_{\mathbf{M}}$. Furthermore, since $\text{Tr}[M_{b|y} F_{\lambda}] \in [0, 1]$ and $\sum_b \text{Tr}[M_{b|y} F_{\lambda}] = 1$, we have that the $p(b|y\lambda)$'s are conditional probabilities.

To show that condition (1) implies condition (0), we can define $p(\lambda|\rho) := \text{Tr}[G_{\lambda}\rho]$, i.e., define a stochastic map that is linear on all quantum states ρ . Thus, $\forall b, y, \rho \in \mathcal{S}(\mathcal{H})$, we have

$$\text{Tr}[M_{b|y}\rho] = \sum_{\lambda} p(b|y\lambda) \text{Tr}[G_{\lambda}\rho], \quad (\text{F7})$$

where $\{p(b|y\lambda)\}_{by}$ respects $\mathcal{O}(\mathbf{M})$, and $p(\lambda|\rho)$ respects $\mathcal{O}_{\mathbf{P}}$. Thus, the statistics of the set of circuits obtained by contracting the multi-measurement with all possible quantum states are classically explainable. Hence, the multi-measurement $\{\{M_{b|y}\}_b\}_y$ is classical. ■

APPENDIX G: PROOF OF COROLLARY 3

We now repeat and then prove Corollary 3.

Corollary 3. If a multi-state consists of a set of states that fits inside a simplex within the quantum state space, then it is classical.

Proof. Assume the existence of such a simplex and denote its vertices by $\{\sigma_{\lambda}\}_{\lambda}$ (by definition, they are affinely independent trace-1 operators, thus linearly independent). As each element of the multi-state $\{\rho_{\cdot|x}\}_x$ is in the convex hull of the $\{\sigma_{\lambda}\}_{\lambda}$, it follows that the span of the former is contained in the span of the latter. Thus, the linearly independent set $\{\sigma_{\lambda}\}_{\lambda}$ forms a basis for the span of the $\{\rho_{\cdot|x}\}$, and so there exists a dual basis $\{H_{\lambda}\}_{\lambda}$ such that [recalling Eq. (A2)]

$$\rho_{\cdot|x} = \sum_{\lambda} \text{Tr}[H_{\lambda} \rho_{\cdot|x}] \sigma_{\lambda}. \quad (\text{G1})$$

Thus, we can define an ontological representation of $\rho_{a|x}$ by

$$p(\lambda|x) := \text{Tr}[H_{\lambda} \rho_{\cdot|x}]. \quad (\text{G2})$$

Given that this is a linear functional on the set of quantum states, it follows that the set $\{p(\lambda|x)\}_x$ satisfies all the operational identities in $\mathcal{O}(\mathbf{P})$ defined by $\{\rho_{\cdot|x}\}_x$. By condition 1 of Theorem 2, one concludes that the multi-state $\{\rho_{\cdot|x}\}_x$ is classical. ■

APPENDIX H: PROOF OF COROLLARY 4

Corollary 4. A multi-state consisting of a set of distinct pure states on a single system is classical if and only if the states are linearly independent.

Proof. The proof for the if direction follows directly from Corollary 2, since a set of linearly independent states forms a simplex. The proof for the only if direction can be shown by contradiction. Suppose the states in the set $\{\rho_{\cdot|x}\}_x$ constituting the multi-state are linearly dependent, then we must have at least one equation of the following form:

$$\sum_x \alpha_x \rho_{\cdot|x} = 0, \quad (\text{H1})$$

where $\alpha_x \neq 0$ for some x . If the set is classical, then there exists a set of states $\{\sigma_\lambda\}_\lambda$ such that $\rho_{\cdot|x}$ can be expressed as a convex mixture of the form $\rho_{\cdot|x} = \sum_\lambda p(\lambda|x) \sigma_\lambda$ where

$$\sum_x \alpha_x p(\lambda|x) = 0. \quad (\text{H2})$$

However, since the $\rho_{\cdot|x}$'s are assumed to be pure, the convex mixture must be trivial; that is, one must have $p(\lambda|x) = \delta_{\lambda,x}$ and $\sigma_x = \rho_{\cdot|x}$. Then, Eq. (H2) reduces to

$$\sum_x \alpha_x \delta_{\lambda,x} = 0, \quad (\text{H3})$$

implying that $\alpha_x = 0$ for all x , contradicting our earlier assumption of linear dependence. Hence $\{\rho_{\cdot|x}\}_x$ must be a linearly independent set. ■

APPENDIX I: PROOF OF PROPOSITION 1

We now repeat and then prove Proposition 1. We note that because the proof is relatively simple, and because the proof of Propositions 4 and 6 are analogous, we do not provide the details of the latter, but rather trust that the reader will see how to generalize the proof of the Proposition 1 to the proofs of Propositions 4 and 6.

Proposition 1. Any circuit with local measurements on a bipartite state [as depicted in Fig. 12(a)] is classically explainable only if the two prepare-measure circuits it induces by boxing to obtain a steering scenario [in one or the other direction, as in Fig. 12(b)] are classically explainable.

Proof. If the statistics generated by a bipartite state $\rho^{A_1 A_2}$ with any multi-measurements $\{\{M_{b_1|y_1}^{A_1}\}_{b_1}\}_{y_1}$ and $\{\{M_{b_2|y_2}^{A_2}\}_{b_2}\}_{y_2}$ are classically explainable, then they can be expressed as

$$p(b_1 b_2 | y_1 y_2) = \sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) p(b_1 | y_1 \lambda_1) p(b_2 | y_2 \lambda_2), \quad (\text{I1})$$

where $p(b_1 | y_1 \lambda_1)$ respects $\mathcal{O}(\mathbf{M}_1)$, $p(b_2 | y_2 \lambda_2)$ respects $\mathcal{O}(\mathbf{M}_2)$, the composite $\sum_{\lambda_1} p(b_1 | y_1 \lambda_1) p(\lambda_1 \lambda_2)$ respects

$\mathcal{O}(\mathbf{M}_1 \circ \mathbf{P})$, the composite $\sum_{\lambda_2} p(b_2 | y_2 \lambda_2) p(\lambda_1 \lambda_2)$ respects $\mathcal{O}(\mathbf{M}_2 \circ \mathbf{P})$, and $p(\lambda_1 \lambda_2)$ respects $\mathcal{O}(\tilde{\mathbf{P}})$ as in Eq. (42) in the main text.

The fact that λ_1 screens off the dependence of b_1 to λ_2 in the bipartite circuit in Fig. 12(a) allows us to define

$$\tilde{p}(b_1 \lambda_2 | y_1) := \sum_{\lambda_1} p(b_1 | y_1 \lambda_1) p(\lambda_1 \lambda_2), \quad (\text{I2})$$

we thus have that

$$p(b_1 b_2 | y_1 y_2) = \sum_{\lambda_2} p(b_2 | y_2 \lambda_2) \tilde{p}(b_1 \lambda_2 | y_1), \quad (\text{I3})$$

which recovers the statistics of a prepare-measure scenario involving the following pair of processes: (1) the preparation $\{\{\tilde{\rho}_{b_1|y_1}^{A_2}\}_{b_1}\}_{y_1}$ on $\mathcal{H}^{A_2}$, where

$$\tilde{\rho}_{b_1|y_1}^{A_2} = \text{Tr}_{A_1} \left[(M_{b_1|y_1}^{A_1} \otimes \mathbb{1}^{A_2}) \rho^{A_1 A_2} \right], \quad (\text{I4})$$

and (2) the measurement $\{\{M_{b_2|y_2}^{A_2}\}_{b_2}\}_{y_2}$ on $\mathcal{H}^{A_2}$. Evidently, $\tilde{p}(b_1 \lambda_2 | y_1)$ respects the identities satisfied by $\{\{\tilde{\rho}_{b_1|y_1}^{A_2}\}_{b_1}\}_{y_1}$, namely, $\mathcal{O}(\mathbf{M}_1 \circ \mathbf{P})$, and $p(b_2 | y_2 \lambda_2)$ respects the identities satisfied by the multi-measurement, namely, $\mathcal{O}(\mathbf{M}_2)$. Thus, such prepare-measure statistics are classically explainable. Using analogous arguments, we infer that the statistics of the prepare-measure scenario on A_1 is also classically explainable. ■

APPENDIX J: PROOF OF FRAME DECOMPOSITION

1. Proof of Theorem 4

Theorem 4. A bipartite state $\rho^{A_1 A_2} \in \mathcal{S}(\mathcal{H}^{A_1} \otimes \mathcal{H}^{A_2})$ is classical if and only if it admits a frame decomposition

$$\rho^{A_1 A_2} = \sum_{\lambda_1, \lambda_2} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}, \quad (\text{J1})$$

where the frame on the operator space V_i is a fixed set of density operators $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i}$, and the dual frame is a set of Hermitian operators $\{H_{\lambda_i}^{A_i}\}_{\lambda_i}$ satisfying $\sum_{\lambda_i} H_{\lambda_i}^{A_i} = \mathcal{P}_{V_i}(\mathbb{1}^{A_i})$ with $\mathcal{P}_{V_i}$ denoting the superoperator projecting onto V_i , and such that

$$\text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \geq 0, \quad \forall \lambda_1, \lambda_2. \quad (\text{J2})$$

Proof. We first prove the only if direction. Assume $\rho^{A_1 A_2}$ is classical in the sense of Definition 7 in the main text. Then, for the statistics generated by the circuits consisting of the bipartite state $\rho^{A_1 A_2}$ on $\mathcal{H}^{A_1 A_2}$ and every possible multi-measurement $\{M_{b_1|y_1}^{A_1}\}_{b_1, y_1}$ on $\mathcal{H}^{A_1}$ and $\{M_{b_2|y_2}^{A_2}\}_{b_2, y_2}$

on $\mathcal{H}^{A_2}$, there exists an ontological model

$$\begin{aligned} & \text{Tr}[(M_{b_1|y_1}^{A_1} \otimes M_{b_2|y_2}^{A_2})\rho^{A_1A_2}] \\ &= \sum_{\lambda_1\lambda_2} p(\lambda_1\lambda_2) p(b_1|y_1\lambda_1) p(b_2|y_2\lambda_2) \end{aligned} \quad (\text{J3a})$$

$$= \sum_{\lambda_1} p(b_1|y_1\lambda_1) \tilde{p}(b_2\lambda_1|y_2) \quad (\text{J3b})$$

$$= \sum_{\lambda_2} p(b_2|y_2\lambda_2) \tilde{p}(b_1\lambda_2|y_1), \quad (\text{J3c})$$

where $\tilde{p}(b_2\lambda_1|y_2)$ and $\tilde{p}(b_1\lambda_2|y_1)$ are defined in Eq. (41) and where the ontological identities in Eq. (42) are satisfied. In particular, these constraints imply that

- (1) From Eq. (42a), for each λ_1 , the map $M_{b_1|y_1}^{A_1} \mapsto p(b_1|y_1\lambda_1)$ is a normalized positive linear functional on the space of all effects on A_1 , i.e., $\mathcal{M}(\mathcal{H}^{A_1})$. The normalization $p(\mathbb{1}^{A_1}|\lambda_1) = 1$ follows from the fact that $\text{Tr}[(\mathbb{1}^{A_1} \otimes \mathbb{1}^{A_2})\rho^{A_1A_2}] = 1$, and by Eq. (J3a), $p(\mathbb{1}^{A_1}|\lambda_i)$ must equal 1. Therefore, by the generalized Gleason theorem, there exists a valid quantum state $\sigma_{\lambda_1}^{A_1}$ such that

$$p(b_1|y_1\lambda_1) = \text{Tr}[M_{b_1|y_1}^{A_1} \sigma_{\lambda_1}^{A_1}]. \quad (\text{J4})$$

- (2) From Eq. (42b), for each λ_2 , the map $M_{b_2|y_2}^{A_2} \mapsto p(b_2|y_2\lambda_2)$ is a normalized positive linear functional on the space of all effects on A_2 , i.e., $\mathcal{M}(\mathcal{H}^{A_2})$. Therefore, by the generalized Gleason theorem, there exists a valid quantum state $\sigma_{\lambda_2}^{A_2}$ such that

$$p(b_2|y_2\lambda_2) = \text{Tr}[M_{b_2|y_2}^{A_2} \sigma_{\lambda_2}^{A_2}]. \quad (\text{J5})$$

- (3) From Eq. (42c), for each λ_2 , the map $\tilde{\rho}_{b_1|y_1}^{A_2} \mapsto \tilde{p}(b_1\lambda_2|y_1)$ [where $\tilde{\rho}_{b_1|y_1}^{A_2}$ is the steered state on A_2 , defined in Eq. (36a)] is a positive linear functional on the operator space $V_2 := \text{Span}(\{\tilde{\rho}_{b_1|y_1}^{A_2}\})$; hence, there exists a Hermitian operator $H_{\lambda_2}^{A_2}$ on $\mathcal{H}^{A_2}$ such that

$$\tilde{p}(b_1\lambda_2|y_1) = \text{Tr}[H_{\lambda_2}^{A_2} \tilde{\rho}_{b_1|y_1}^{A_2}]. \quad (\text{J6})$$

- (4) From Eq. (42d), for each λ_1 the map $\tilde{\rho}_{b_2|y_2}^{A_1} \mapsto \tilde{p}(b_2\lambda_1|y_2)$ [where $\tilde{\rho}_{b_2|y_2}^{A_1}$ is the steered state on A_1 , defined in Eq. (36b)] is a positive linear functional on the operator space $V_1 = \text{Span}(\{\tilde{\rho}_{b_2|y_2}^{A_1}\})$; hence, there exists a Hermitian operator $H_{\lambda_1}^{A_1}$ on $\mathcal{H}^{A_1}$ such that

$$\tilde{p}(b_2\lambda_1|y_2) = \text{Tr}[H_{\lambda_1}^{A_1} \tilde{\rho}_{b_2|y_2}^{A_1}]. \quad (\text{J7})$$

Plugging Eqs. (J4) and (J7) into Eq. (J3b) yields

$$\begin{aligned} & \text{Tr}[(M_{b_1|y_1}^{A_1} \otimes M_{b_2|y_2}^{A_2})\rho^{A_1A_2}] \\ &= \sum_{\lambda_1} \text{Tr}[M_{b_1|y_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \text{Tr}[H_{\lambda_1}^{A_1} \tilde{\rho}_{b_2|y_2}^{A_1}]. \end{aligned} \quad (\text{J8})$$

Recalling the definition of $\tilde{\rho}_{b_2|y_2}^{A_1}$ provided in Eq. (36b), we have

$$\text{Tr}[(M_{b_1|y_1}^{A_1} \otimes M_{b_2|y_2}^{A_2})\rho^{A_1A_2}] = \text{Tr}[M_{b_1|y_1}^{A_1} \tilde{\rho}_{b_2|y_2}^{A_1}], \quad (\text{J9})$$

and combining these two expressions, we obtain

$$\text{Tr}[M_{b_1|y_1}^{A_1} \tilde{\rho}_{b_2|y_2}^{A_1}] = \sum_{\lambda_1} \text{Tr}[M_{b_1|y_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \text{Tr}[H_{\lambda_1}^{A_1} \tilde{\rho}_{b_2|y_2}^{A_1}].$$

Since this holds for all $M_{b_1|y_1}^{A_1} \in \mathcal{M}(\mathcal{H}^{A_1})$ and these are tomographically complete, it follows that

$$\tilde{\rho}_{b_2|y_2}^{A_1} = \sum_{\lambda_1} \text{Tr}[\tilde{\rho}_{b_2|y_2}^{A_1} H_{\lambda_1}^{A_1}] \sigma_{\lambda_1}^{A_1}. \quad (\text{J10})$$

That is, $\{\sigma_{\lambda_1}^{A_1}\}_{\lambda_1}$ and $\{H_{\lambda_1}^{A_1}\}_{\lambda_1}$ define a frame and its dual frame on $V_1 = \text{Span}(\{\tilde{\rho}_{b_2|y_2}^{A_1}\})$. Moreover, taking the trace of both sides yields

$$\text{Tr}[\tilde{\rho}_{b_2|y_2}^{A_1}] = \text{Tr}[\tilde{\rho}_{b_2|y_2}^{A_1} \sum_{\lambda_1} H_{\lambda_1}^{A_1}]. \quad (\text{J11})$$

Noting that this holds for all $\tilde{\rho}_{b_2|y_2}^{A_1} \in V_1$ implies that $\sum_{\lambda_1} H_{\lambda_1}^{A_1} = \mathcal{P}_{V_1}(\mathbb{1}^{A_1})$, where we denote the superoperator that projects onto the operator space V_1 by $\mathcal{P}_{V_1}$.

Similarly, substituting Eqs. (J5) and (J6) into Eq. (J3c), we can prove that $\{\sigma_{\lambda_2}^{A_2}\}_{\lambda_2}$, and the set of Hermitian operators $\{H_{\lambda_2}^{A_2}\}_{\lambda_2}$ [with $\sum_{\lambda_2} H_{\lambda_2}^{A_2} = \mathcal{P}_{V_2}(\mathbb{1}^{A_2})$] defines a pair of frame and dual frame on $V_2 = \text{Span}(\{\tilde{\rho}_{b_1|y_1}^{A_2}\})$, and in particular,

$$\tilde{\rho}_{b_1|y_1}^{A_2} = \sum_{\lambda_2} \text{Tr}[\tilde{\rho}_{b_1|y_1}^{A_2} H_{\lambda_2}^{A_2}] \sigma_{\lambda_2}^{A_2}. \quad (\text{J12})$$

- (5) Finally, using the ontological identities in Eq. (42e) and substituting Eqs. (J6) and (J7), we obtain

$$\begin{aligned}
p(\lambda_1 \lambda_2) &= \sum_{b_1 b_2 y_1 y_2} \alpha_{b_1 b_2 y_1 y_2} \tilde{p}(b_1 \lambda_2 | y_1) \tilde{p}(b_2 \lambda_1 | y_2) \\
&= \sum_{b_1 b_2 y_1 y_2} \alpha_{b_1 b_2 y_1 y_2} \text{Tr}[H_{\lambda_2}^{A_2} \tilde{\rho}_{b_1 | y_1}^{A_2}] \\
&\quad \times \text{Tr}[H_{\lambda_1}^{A_1} \tilde{\rho}_{b_2 | y_2}^{A_1}]. \tag{J13}
\end{aligned}$$

Recalling the operational identity in Eq. (38), we thus have

$$p(\lambda_1 \lambda_2) = \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}]. \tag{J14}$$

Combining it with Eqs. (J3a), (J4)(J5) gives

$$\begin{aligned}
&\text{Tr}[(M_{b_1 | y_1}^{A_1} \otimes M_{b_2 | y_2}^{A_2}) \rho^{A_1 A_2}] \\
&= \sum_{\lambda_1 \lambda_2} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \text{Tr}[M_{b_1 | y_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \\
&\quad \times \text{Tr}[M_{b_2 | y_2}^{A_2} \sigma_{\lambda_2}^{A_2}]. \tag{J15}
\end{aligned}$$

Since this holds for all $M_{b_i | y_i}^{A_i} \in \mathcal{M}(\mathcal{H}^{A_i})$, which are tomographically complete, we conclude that

$$\rho^{A_1 A_2} = \sum_{\lambda_1 \lambda_2} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \sigma_{\lambda_1}^{A_1} \otimes \sigma_{\lambda_2}^{A_2}. \tag{J16}$$

Moreover, the constraint $\text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \geq 0$ in Eq. (J2) holds by Eq. (J14) and the existence of the ontological model in Eq. (J3a).

Now we prove the if direction. Suppose there exists a frame decomposition of $\rho^{A_1 A_2}$ as in Eq. (J1) and satisfying the requisite constraints. For the bipartite state $\rho^{A_1 A_2}$, any effect $M_{b_1 | y_1}^{A_1} \in \mathcal{M}(\mathcal{H}^{A_1})$ and any effect $M_{b_2 | y_2}^{A_2} \in \mathcal{M}(\mathcal{H}^{A_2})$, we can define ontological representations of each of these via the linear functionals induced by the frames and dual frames:

$$p(\lambda_1 \lambda_2) := \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}], \tag{J17a}$$

$$p(b_1 | y_1 \lambda_1) = \text{Tr}[\sigma_{\lambda_1}^{A_1} M_{b_1 | y_1}^{A_1}], \tag{J17b}$$

$$p(b_2 | y_2 \lambda_2) = \text{Tr}[\sigma_{\lambda_2}^{A_2} M_{b_2 | y_2}^{A_2}]. \tag{J17c}$$

The non-negativity and normalization of these ontological representations follow from Eq. (J2) together with the fact that the frame $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i}$ is a set of density operators.

Next, we show that the effective ontological representations obtained by sequential composition depend linearly on the corresponding effective quantum processes. This follows from the fact that $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i}$ and $\{H_{\lambda_i}^{A_i}\}_{\lambda_i}$ form a frame and its dual frame on V_i . In particular,

$$\begin{aligned}
\tilde{p}(b_1 \lambda_2 | y_1) &:= \sum_{\lambda_1} p(\lambda_1 \lambda_2) p(b_1 | y_1 \lambda_1) \\
&= \sum_{\lambda_1} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \text{Tr}[M_{b_1 | y_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \\
&= \text{Tr}[(M_{b_1 | y_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] = \text{Tr}[H_{\lambda_2}^{A_2} \tilde{\rho}_{b_1 | y_1}^{A_2}], \tag{J17d}
\end{aligned}$$

where we use Eqs. (J17a) and (J17b) in the first equality, and the definition of $\tilde{\rho}_{b_1 | y_1}^{A_2}$, provided in Eq. (36a), in the last equality. Similarly, we can show that

$$\begin{aligned}
\tilde{p}(b_2 \lambda_1 | y_2) &:= \sum_{\lambda_2} p(\lambda_1 \lambda_2) p(b_2 | y_2 \lambda_2) \\
&= \sum_{\lambda_2} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \text{Tr}[M_{b_2 | y_2}^{A_2} \sigma_{\lambda_2}^{A_2}] \\
&= \text{Tr}[(H_{\lambda_1}^{A_1} \otimes M_{b_2 | y_2}^{A_2}) \rho^{A_1 A_2}] = \text{Tr}[H_{\lambda_1}^{A_1} \tilde{\rho}_{b_2 | y_2}^{A_1}] \tag{J17e}
\end{aligned}$$

using Eqs. (J17a) and (J17c) in the first equality, and the definition of $\tilde{\rho}_{b_2 | y_2}^{A_1}$, provided in Eq. (36b), in the last equality.

Finally, the ontological model reproduces the empirical statistics

$$\begin{aligned}
&\sum_{\lambda_1 \lambda_2} p(\lambda_1 \lambda_2) p(b_1 | y_1 \lambda_1) p(b_2 | y_2 \lambda_2) \\
&= \sum_{\lambda_1 \lambda_2} \text{Tr}[(H_{\lambda_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \text{Tr}[M_{b_1 | y_1}^{A_1} \sigma_{\lambda_1}^{A_1}] \text{Tr}[M_{b_2 | y_2}^{A_2} \sigma_{\lambda_2}^{A_2}] \\
&= \sum_{\lambda_2} \text{Tr}[(M_{b_1 | y_1}^{A_1} \otimes H_{\lambda_2}^{A_2}) \rho^{A_1 A_2}] \text{Tr}[M_{b_2 | y_2}^{A_2} \sigma_{\lambda_2}^{A_2}] \\
&= \text{Tr}[(M_{b_1 | y_1}^{A_1} \otimes M_{b_2 | y_2}^{A_2}) \rho^{A_1 A_2}], \tag{J18}
\end{aligned}$$

where in the last two steps, we again use the properties of frame representations.

Since all these ontological representations depend linearly on the corresponding physical processes and are non-negative, the circuit is classically explainable. Moreover, since we allowed all possible factorizing dual processes, namely, all local multi-measurements, it follows that the bipartite state $\rho^{A_1 A_2}$ is classical by Definition 7. ■

2. Proof of Theorem 5

Theorem 5. A channel $\mathcal{E}^{B|A}: \mathcal{L}(\mathcal{H}^A) \mapsto \mathcal{L}(\mathcal{H}^B)$ is classical if and only if it admits a frame representation

$$\mathcal{E}^{B|A}(\cdot) = \sum_{\lambda_A \lambda_B} \sigma_{\lambda_B}^B \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A(\cdot)], \tag{J19}$$

where the frame on V_A is a set of Hermitian operators $\{F_{\lambda_A}^A\}_{\lambda_A}$ each with unit trace, $\text{Tr}[F_{\lambda_A}^A] = 1$, and the dual

frame on V_A is a POVM $\{G_{\lambda_A}^A\}_{\lambda_A}$ satisfying $\sum_{\lambda_A} G_{\lambda_A}^A = \mathcal{P}_{V_A}(\mathbb{1}^A) = \mathbb{1}^A$, while the frame on V_B is a set of density operators $\{\sigma_{\lambda_B}^B\}_{\lambda_B}$ and the dual frame on V_B is a set of Hermitian operators $\{H_{\lambda_B}^B\}_{\lambda_B}$ satisfying $\sum_{\lambda_B} H_{\lambda_B}^B = \mathcal{P}_{V_B}(\mathbb{1}^B)$, with $\mathcal{P}_{V_{B(A)}}$ denoting the superoperator projecting onto $V_{B(A)}$, and such that

$$\text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \geq 0, \quad \forall \lambda_A, \lambda_B. \quad (\text{J20})$$

Proof. We first prove the only if direction. Assume $\mathcal{E}^{B|A}$ is classical in the sense of Definition 9 in the main text. Then, for the statistics generated in the prepare-transform-measure circuits consisting of the channel $\mathcal{E}^{B|A}$ and every possible multi-source $\{\{p(a|x)\rho_{a|x}^A\}_x\}$ and multi-measurement $\{\{M_{b|y}^B\}_y\}$, there exists an ontological model

$$\begin{aligned} & \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] \\ &= \sum_{\lambda_A \lambda_B} p(b|y\lambda_B) p(\lambda_B|\lambda_A) p(a\lambda_A|x) \end{aligned} \quad (\text{J21a})$$

$$= \sum_{\lambda_B} p(b|y\lambda_B) \tilde{p}(a\lambda_B|x) \quad (\text{J21b})$$

$$= \sum_{\lambda_A} \tilde{p}(b|y\lambda_A) p(a\lambda_A|x), \quad (\text{J21c})$$

where $\tilde{p}(a\lambda_B|x)$ and $\tilde{p}(b|y\lambda_A)$ are defined in Eq. (68), and where the ontological identities in Eq. (69) are respected. In particular, these constraints imply that

- (1) From Eq. (69a), for each λ_B , the map $M_{b|y}^B \mapsto p(b|y\lambda_B)$ is a normalized positive linear functional on the space of all effects on B , i.e., $\mathcal{M}(\mathcal{H}^B)$. Normalization $p(\mathbb{1}^B|\lambda_B) = 1$ follows from $\text{Tr}[\mathbb{1}^B \mathcal{E}^{B|A}(\rho^A)] = 1$, and by Eq. (J21a), $p(\mathbb{1}^B|\lambda_B)$ must equal 1. Thus, by the generalized Gleason theorem, there exists a quantum state $\sigma_{\lambda_B}^B$ such that

$$p(b|y\lambda_B) = \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B]. \quad (\text{J22})$$

- (2) From Eq. (69b), for each λ_A , the map $p(a|x)\rho_{a|x}^A \mapsto p(a\lambda_A|x)$ is a positive linear functional on the space of all unnormalized states on A . Therefore, by the Riesz representation theorem, there exists a positive semidefinite operator $G_{\lambda_A}^A$ such that

$$p(a\lambda_A|x) = \text{Tr}[G_{\lambda_A}^A p(a|x)\rho_{a|x}^A]. \quad (\text{J23})$$

- (3) From Eq. (69c), for each λ_A , the map $\tilde{M}_{b|y}^A \mapsto \tilde{p}(b|y\lambda_A)$ is a normalized positive linear functional on $V_A = \text{Span}(\{\tilde{M}_{b|y}^A\})$ with $\tilde{M}_{b|y}^A = [\mathcal{E}^{B|A}]^\dagger(M_{b|y}^B)$; therefore, there exists a Hermitian operator $F_{\lambda_A}^A$

such that

$$\tilde{p}(b|y\lambda_A) = \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A]. \quad (\text{J24})$$

Since $\sum_b \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A] = 1$, $F_{\lambda_A}^A$ is trace-1.

- (4) From Eq. (69d), for each λ_B , the map $\tilde{\rho}_{a|x}^B \mapsto \tilde{p}(a\lambda_B|x)$ is a positive linear functional on $V_B = \text{Span}(\{\tilde{\rho}_{a|x}^B\})$ with $\tilde{\rho}_{a|x}^B = \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)$. Thus, there exists a Hermitian operator $H_{\lambda_B}^B$ such that

$$\tilde{p}(a\lambda_B|x) = \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B]. \quad (\text{J25})$$

Substituting Eq. (J22), (J25) into Eq. (J21b) yields

$$\begin{aligned} & \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] \\ &= \sum_{\lambda_B} \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B] \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B]. \end{aligned} \quad (\text{J26})$$

Recalling the definition of $\tilde{\rho}_{a|x}^B$ in Eq. (62b), we have

$$\text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] := \text{Tr}[M_{b|y}^B \tilde{\rho}_{a|x}^B], \quad (\text{J27})$$

and combining these two expressions, we obtain

$$\text{Tr}[M_{b|y}^B \tilde{\rho}_{a|x}^B] = \sum_{\lambda_B} \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B] \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B]. \quad (\text{J28})$$

Importantly, since this holds for all $M_{b|y}^B \in \mathcal{M}(\mathcal{H}^B)$, which are tomographically complete, this implies

$$\tilde{\rho}_{a|x}^B = \sum_{\lambda_B} \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B] \sigma_{\lambda_B}^B. \quad (\text{J29})$$

That is, $\{\sigma_{\lambda_B}^B\}_{\lambda_B}$ and $\{H_{\lambda_B}^B\}_{\lambda_B}$ define a frame and its dual frame on $V_B = \text{Span}(\{\tilde{\rho}_{a|x}^B\})$. Moreover, taking the trace of both sides yields

$$\text{Tr}[\tilde{\rho}_{a|x}^B] = \sum_{\lambda_B} \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B]. \quad (\text{J30})$$

Noting that this holds for all $\tilde{\rho}_{a|x}^B$ on V_B implies that $\sum_{\lambda_B} H_{\lambda_B}^B = \mathcal{P}_{V_B}(\mathbb{1}^B)$, where we denote the superoperator that projects onto V_B by $\mathcal{P}_{V_B}$.

Similarly, substituting Eqs. (J23) and (J24) into Eq. (J21c), we have

$$\begin{aligned} & \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] \\ &= \sum_{\lambda_A} \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A] \text{Tr}[G_{\lambda_A}^A p(a|x)\rho_{a|x}^A]. \end{aligned} \quad (\text{J31})$$

Recalling the definition of $\tilde{M}_{b|y}^A$ in Eq. (62b), we have

$$\text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] = \text{Tr}[\tilde{M}_{b|y}^A p(a|x)\rho_{a|x}^A], \quad (\text{J32})$$

and combining these two expressions, we obtain

$$\begin{aligned} & \text{Tr}[\tilde{M}_{b|y}^A p(a|x)\rho_{a|x}^A] \\ &= \sum_{\lambda_A} \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A] \text{Tr}[G_{\lambda_A}^A p(a|x)\rho_{a|x}^A]. \end{aligned} \quad (\text{J33})$$

Importantly, since this holds for all unnormalized states $p(a|x)\rho_{a|x}^A$ with $\rho_{a|x}^A \in \mathcal{S}(\mathcal{H}^A)$, which are tomographically complete, we have

$$\tilde{M}_{b|y}^A = \sum_{\lambda_A} \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A] G_{\lambda_A}^A. \quad (\text{J34})$$

That is, $\{F_{\lambda_A}^A\}_{\lambda_A}$ and $\{G_{\lambda_A}^A\}_{\lambda_A}$ define a frame and dual frame on $V_A = \text{Span}(\{\tilde{M}_{b|y}^A\})$. Moreover, taking the summation over b on both sides yields

$$\begin{aligned} \sum_b \tilde{M}_{b|y}^A &= \sum_b \sum_{\lambda_A} \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A] G_{\lambda_A}^A \\ &= \sum_{\lambda_A} G_{\lambda_A}^A = \mathcal{P}_V(\mathbb{1}^A) = \mathbb{1}^A. \end{aligned} \quad (\text{J35})$$

Therefore, $\{G_{\lambda_A}^A\}_{\lambda_A}$ is a valid POVM.

- (5) Finally, plugging Eqs. (J24) and (J25) into the assumed decomposition in Eq. (69e), we obtain

$$\begin{aligned} p(\lambda_B|\lambda_A) &= \sum_{abxy} \alpha_{abxy} \tilde{p}(b|y\lambda_A) \tilde{p}(a\lambda_B|x) \\ &= \sum_{abxy} \alpha_{abxy} \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A] \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B]. \end{aligned} \quad (\text{J36})$$

Recalling the operational identity in Eq. (65), we thus have

$$p(\lambda_B|\lambda_A) = \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)]. \quad (\text{J37})$$

Combining it with Eq. (J22), (J23), and (J21) yields

$$\begin{aligned} & \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] \\ &= \sum_{\lambda_A \lambda_B} \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B] \\ &\quad \times \text{Tr}[G_{\lambda_A}^A p(a|x)\rho_{a|x}^A], \end{aligned} \quad (\text{J38})$$

since this holds for all $M_{b|y}^B$ and $p(a|x)\rho_{a|x}^A$, which are tomographically complete, we have

$$\mathcal{E}^{B|A}(\cdot) = \sum_{\lambda_A \lambda_B} \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A(\cdot)] \sigma_{\lambda_B}^B. \quad (\text{J39})$$

Now let us prove the if part. Suppose there exists a frame decomposition of the channel of $\mathcal{E}^{B|A}$ as in Eq. (J19), then for the channel $\mathcal{E}^{B|A}$, any effect $M_{b|y}^B$ and any unnormalized state $p(a|x)\rho_{a|x}^A$, we can define ontological representations of each of these processes via the linear functionals induced by the frames and dual frames:

$$p(\lambda_B|\lambda_A) := \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)], \quad (\text{J40a})$$

$$p(b|y\lambda_B) := \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B], \quad (\text{J40b})$$

$$p(a\lambda_A|x) := \text{Tr}[G_{\lambda_A}^A p(a|x)\rho_{a|x}^A], \quad (\text{J40c})$$

where the non-negativity and normalization are ensured by Eq. (J20), $\sigma_{\lambda_B}^B$ being a density operator, and $\{G_{\lambda_A}^A\}_{\lambda_A}$ being a POVM.

We can also show that the effective ontological processes obtained by sequential compositions are also linear to the corresponding effective quantum processes obtained by sequential compositions. This is because $\{\sigma_{\lambda_i}^{A_i}\}_{\lambda_i}$ and $\{H_{\lambda_i}^{A_i}\}_{\lambda_i}$ are a frame and its dual frame on V_i . In particular,

$$\begin{aligned} \tilde{p}(a\lambda_B|x) &:= \sum_{\lambda_A} p(\lambda_B|\lambda_A) p(a\lambda_A|x) \\ &= \sum_{\lambda_A} \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A p(a|x)\rho_{a|x}^A] \\ &= \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(p(a|x)\rho_{a|x}^A)] = \text{Tr}[H_{\lambda_B}^B \tilde{\rho}_{a|x}^B], \end{aligned} \quad (\text{J40d})$$

where we use Eqs. (J40a) and (J40c) in the first equality, and the definition of $\tilde{\rho}_{a|x}^B$ in the last equality. Similarly, using Eqs. (J40a) and (J40b), we can show that

$$\begin{aligned} \tilde{p}(b|y\lambda_A) &:= \sum_{\lambda_B} p(b|y\lambda_B) p(\lambda_B|\lambda_A) \\ &= \sum_{\lambda_B} \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B] \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \\ &= \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] = \text{Tr}[\tilde{M}_{b|y}^A F_{\lambda_A}^A]. \end{aligned} \quad (\text{J40e})$$

Finally, the ontological model also reproduces the empirical statistics:

$$\begin{aligned}
 & \sum_{\lambda_A \lambda_B} p(b|y\lambda_B) p(\lambda_B|\lambda_A) p(a\lambda_A|x) \\
 &= \sum_{\lambda_A \lambda_B} \text{Tr}[M_{b|y}^B \sigma_{\lambda_B}^B] \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A p(a|x) \rho_{a|x}^A] \\
 &= \sum_{\lambda_A} \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A p(a|x) \rho_{a|x}^A] \\
 &= \text{Tr}[M_{b|y}^B \mathcal{E}^{B|A}(p(a|x) \rho_{a|x}^A)]. \quad (\text{J41})
 \end{aligned}$$

In the last two equations, we again use properties of frame representations.

Since all these ontological representations depend linearly on the corresponding physical processes and are non-negative, the circuit is classically explainable. Moreover, since we allowed all possible factorizing dual processes, this, in turn, implies that the channel $\mathcal{E}^{B|A}$ is classical by Definition 9. ■

3. Proof of Theorem 6

Theorem 6. A multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$, where $\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}: \mathcal{L}(\otimes_{i=1}^n \mathcal{H}^{A_i}) \mapsto \mathcal{L}(\otimes_{j=1}^m \mathcal{H}^{B_j})$, is classical if and only if it admits a frame representation

$$\begin{aligned}
 \mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\cdot) &= \sum_{\lambda_{\vec{A}} \lambda_{\vec{B}}} \bigotimes_j \sigma_{\lambda_{B_j}}^{B_j} \text{Tr} \left[\left(\bigotimes_j H_{\lambda_{B_j}}^{B_j} \right) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\bigotimes_i F_{\lambda_{A_i}}^{A_i} \right) \right] \\
 &\quad \times \text{Tr} \left[\left(\bigotimes_i G_{\lambda_{A_i}}^{A_i} \right) (\cdot) \right], \quad (\text{J42})
 \end{aligned}$$

where the frame on the operator space V_{B_j} is a fixed set of density operators $\{\sigma_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$, and the dual frame is a set of Hermitian operators $\{H_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$ satisfying $\sum_{\lambda_{B_j}} H_{\lambda_{B_j}}^{B_j} = \mathcal{P}_{V_{B_j}}(\mathbb{1}^{B_j})$; the frame on the operator space V_{A_i} is a set of Hermitian operators $\{F_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ satisfying $\text{Tr}[F_{\lambda_{A_i}}^{A_i}] = 1$, and the dual frame is a POVM $\{G_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ satisfying $\sum_{\lambda_{A_i}} G_{\lambda_{A_i}}^{A_i} = \mathcal{P}_{V_{A_i}}(\mathbb{1}^{A_i}) = \mathbb{1}^{A_i}$, and such that

$$\text{Tr} \left[\left(\bigotimes_j H_{\lambda_{B_j}}^{B_j} \right) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\bigotimes_i F_{\lambda_{A_i}}^{A_i} \right) \right] \geq 0, \quad \forall \lambda_{\vec{A}}, \lambda_{\vec{B}}, c, z. \quad (\text{J43})$$

Proof. We first prove the only if direction. Assume the multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ is classical. Then, for the statistics generated in the multipartite prepare-transform-measure circuit consisting of the multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ and every possible multi-source $\{\{p(a_i|x_i) \rho_{a_i|x_i}^{A_i}\}_{a_i}\}_{x_i}$ on $\mathcal{H}^{A_i}$ and multi-measurement $\{\{M_{b_j|y_j}^{B_j}\}_{b_j}\}_{y_j}$ on $\mathcal{H}^{B_j}$, there exists an

ontological model

$$\begin{aligned}
 & \text{Tr} \left[\left(\bigotimes_{j \in [m]} M_{b_j|y_j}^{B_j} \right) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\bigotimes_{i \in [n]} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right) \right] \\
 &= \sum_{\lambda_{\vec{B}} \lambda_{\vec{A}}} p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) \prod_j p(b_j|y_j\lambda_{B_j}) \prod_i p(a_i\lambda_{A_i}|x_i) \quad (\text{J44a})
 \end{aligned}$$

$$= \sum_{\lambda_{B_j}} p(b_j|y_j\lambda_{B_j}) \tilde{p}(\vec{a}\vec{b}_j c \lambda_{B_j} | \vec{x}\vec{y}_j z) \quad (\text{J44b})$$

$$= \sum_{\lambda_{A_i}} p(a_i\lambda_{A_i}|x_i) \tilde{p}(\vec{a}_i \vec{b} c | \vec{x}_i \vec{y} z \lambda_{A_i}), \quad (\text{J44c})$$

where the ontological identities in Eq. (83) are respected. Here we denote $\vec{j} = \{j' | j' \neq j\}$ and $\vec{i} = \{i' | i' \neq i\}$ and define

$$\begin{aligned}
 \tilde{p}(\vec{a}\vec{b}_j c \lambda_{B_j} | \vec{x}\vec{y}_j z) &= \sum_{\lambda_{\vec{B}} \lambda_{\vec{A}}} p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) \prod_{j' \neq j} p(b_{j'}|y_{j'}\lambda_{B_{j'}}) \\
 &\quad \times \prod_i p(a_i\lambda_{A_i}|x_i) \quad (\text{J45a})
 \end{aligned}$$

$$\begin{aligned}
 \tilde{p}(\vec{a}_i \vec{b} c | \vec{x}_i \vec{y} z \lambda_{A_i}) &= \sum_{\lambda_{\vec{B}} \lambda_{\vec{A}}} p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) \prod_j p(b_j|y_j\lambda_{B_j}) \\
 &\quad \times \prod_{i' \neq i} p(a_{i'}\lambda_{A_{i'}}|x_{i'}). \quad (\text{J45b})
 \end{aligned}$$

Specifically, the constraints in Eq. (83) then imply that

- (1) From Eq. (83a), for each $j \in [m]$ each λ_{B_j} , the map $M_{b_j|y_j}^{B_j} \mapsto p(b_j|y_j, \lambda_{B_j})$ is a normalized positive linear functional on the space of all effects on B_j , i.e., $\mathcal{M}(\mathcal{H}^{B_j})$. Thus, by the generalized Gleason theorem, there exists a quantum state $\sigma_{\lambda_{B_j}}^{B_j}$ such that

$$p(b_j|y_j\lambda_{B_j}) = \text{Tr} \left[M_{b_j|y_j}^{B_j} \sigma_{\lambda_{B_j}}^{B_j} \right]. \quad (\text{J46})$$

- (22) From Eq. (83b), for each $i \in [n]$, for each λ_{A_i} , the map $p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \mapsto p(a_i\lambda_{A_i}|x_i)$ is a positive linear functional on the space of all unnormalized states. Therefore, there exists a positive semidefinite operator $G_{\lambda_{A_i}}^{A_i}$ such that

$$p(a_i\lambda_{A_i}|x_i) = \text{Tr} \left[G_{\lambda_{A_i}}^{A_i} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right]. \quad (\text{J47})$$

There are many ontological identities included in Eq. (83c); however, it turns out that it is sufficient to consider those for the ontological representation in Eq. (J45) defined for the effective

physical processes

$$\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} = \text{Tr}_{\vec{B}_j} \left[M_{\vec{b}|\vec{y}_j}^{\vec{B}} \mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\tilde{\rho}_{\vec{a}|\vec{x}}^{\vec{A}}) \right], \quad (\text{J48a})$$

$$\tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i} = \text{Tr}_{\vec{A}_i} \left[\tilde{\rho}_{\vec{a}|\vec{x}_i}^{\vec{A}_i} [\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}]^\dagger (M_{\vec{b}|\vec{y}}^{\vec{B}}) \right], \quad (\text{J48b})$$

where we denote $M_{\vec{b}|\vec{y}}^{\vec{B}} = \otimes_j M_{b_j|y_j}^{B_j}$ and $M_{\vec{b}|\vec{y}}^{\vec{B}} := \mathbb{1}^{B_j} \otimes_{j' \neq j} M_{b_{j'}|y_{j'}}^{B_{j'}}$, similarly, $\tilde{\rho}_{\vec{a}|\vec{x}}^{\vec{A}} := \otimes_i p(a_i|x_i) \rho_{a_i|x_i}^{A_i}$ and $\tilde{\rho}_{\vec{a}|\vec{x}_i}^{\vec{A}_i} := \mathbb{1}^{A_i} \otimes_{i' \neq i} p(a_{i'}|x_{i'}) \rho_{a_{i'}|x_{i'}}^{A_{i'}}$.

- (3) As a special case of Eq. (83c), for each $j \in [m]$ and each λ_{B_j} , the map $\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \mapsto \tilde{p}(\vec{a}\vec{b}c|\lambda_{B_j}|\vec{x}\vec{y}z)$ is a positive linear functional on the space of all states $\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \in V_{B_j}$ in Eq. (J48a). Thus, there exists a Hermitian operator $H_{\lambda_{B_j}}^{B_j}$ such that

$$\tilde{p}(\vec{a}\vec{b}c|\lambda_{B_j}|\vec{x}\vec{y}z) = \text{Tr} \left[H_{\lambda_{B_j}}^{B_j} \tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right]. \quad (\text{J49})$$

- (4) As another special case of Eq. (83c), for each $i \in [n]$ and each λ_{A_i} , the map $\tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i} \mapsto \tilde{p}(\vec{a}\vec{b}c|\lambda_{A_i}|\vec{x}\vec{y}z)$ is a positive linear functional on the space of all effects $\tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i} \in V_{A_i}$ in Eq. (48b). Thus, by the Riesz representation theorem, there exists a Hermitian operator $F_{\lambda_{A_i}}^{A_i}$ such that

$$\tilde{p}(\vec{a}\vec{b}c|\lambda_{A_i}|\vec{x}\vec{y}z) = \text{Tr} \left[F_{\lambda_{A_i}}^{A_i} \tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i} \right]. \quad (\text{J50})$$

Plugging Eqs. (J46) and (J49) into Eq. (J44b) yields

$$\begin{aligned} & \text{Tr} \left[\left(\otimes_{j \in [m]} M_{b_j|y_j}^{B_j} \right) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\otimes_{i \in [n]} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right) \right] \\ &= \sum_{\lambda_{B_j}} \text{Tr} \left[M_{b_j|y_j}^{B_j} \sigma_{\lambda_{B_j}}^{B_j} \right] \text{Tr} \left[H_{\lambda_{B_j}}^{B_j} \tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right]. \end{aligned} \quad (\text{J51})$$

Combining it with Eq. (J48a), we have

$$\begin{aligned} & \text{Tr} \left[M_{b_j|y_j}^{B_j} \tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right] \\ &= \sum_{\lambda_{B_j}} \text{Tr} \left[M_{b_j|y_j}^{B_j} \sigma_{\lambda_{B_j}}^{B_j} \right] \text{Tr} \left[H_{\lambda_{B_j}}^{B_j} \tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right]. \end{aligned} \quad (\text{J52})$$

Importantly, since this holds for all $M_{b_j|y_j}^{B_j} \in \mathcal{M}(\mathcal{H}_{B_j})$, which are tomographically complete, this implies

$$\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} = \sum_{\lambda_{B_j}} \text{Tr} \left[H_{\lambda_{B_j}}^{B_j} \tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right] \sigma_{\lambda_{B_j}}^{B_j}. \quad (\text{J53})$$

That is, $\{\sigma_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$ and $\{H_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$ define a frame and its dual frame on $V_{B_j} = \text{Span}(\{\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j}\})$. Moreover, taking the trace of both sides yields

$$\text{Tr} \left[\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right] = \sum_{\lambda_{B_j}} \text{Tr} \left[H_{\lambda_{B_j}}^{B_j} \tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \right]. \quad (\text{J54})$$

Noting that this holds for all $\tilde{\rho}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{B_j} \in V_{B_j}$, we thus have $\sum_{\lambda_{B_j}} H_{\lambda_{B_j}}^{B_j} = \mathcal{P}_{V_{B_j}}(\mathbb{1}^{B_j})$, where $\mathcal{P}_{V_{B_j}}$ is the superoperator that projects onto the operator space V_{B_j} .

Similarly, by plugging Eqs. (J47) and (J50) into Eq. (J44c), we can prove that $\{F_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ and a POVM $\{G_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ defines a frame and dual frame on $V_{A_i} = \text{Span}(\{\tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i}\})$, and in particular,

$$\tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i} = \sum_{\lambda_{A_i}} \text{Tr} \left[F_{\lambda_{A_i}}^{A_i} \tilde{M}_{\vec{a}\vec{b}c|\vec{x}\vec{y}z}^{A_i} \right] G_{\lambda_{A_i}}^{A_i}. \quad (\text{J55})$$

- (5) Finally, we now consider the ontological identities in Eq. (83d). Again, there are many ontological identities therein. However, it turns out that it is sufficient to consider some special ones. Specifically, for any fixed c and z , we consider operational identity of the form of

$$\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\cdot) = \sum_{\substack{\{c^i \vec{b}^i \vec{y}^i\}_i \{c^j \vec{b}^j \vec{y}^j\}_j \\ \{c^i \vec{a}^i \vec{x}^i\}_i \{c^j \vec{a}^j \vec{x}^j\}_j}} \alpha_{\{c^i \vec{b}^i \vec{y}^i\}_i \{c^j \vec{b}^j \vec{y}^j\}_j}^{\{c^i \vec{a}^i \vec{x}^i\}_i \{c^j \vec{a}^j \vec{x}^j\}_j} \text{Tr} \left[\left(\otimes_i \tilde{M}_{\vec{a}^i \vec{b}^i c|\vec{x}^i \vec{y}^i}^{A_i} \right) (\cdot) \right] \otimes_j \tilde{\rho}_{\vec{a}^j \vec{b}^j c|\vec{x}^j \vec{y}^j}^{B_j}, \quad (\text{J56})$$

where the superscripts i and j are indices to distinguish different settings and outcome variables. The associated ontological identities are then given as

$$\begin{aligned}
p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) &= \sum_{\substack{\{c^i\vec{b}^i\vec{y}^i\}_i\{\vec{c}^j\vec{b}^j\vec{y}^j\}_j \\ \{z^i\vec{a}^i\vec{x}^i\}_i\{z^j\vec{a}^j\vec{x}^j\}_j}} \alpha_{\{c^i\vec{b}^i\vec{y}^i\}_i\{\vec{c}^j\vec{b}^j\vec{y}^j\}_j}^{\{z^i\vec{a}^i\vec{x}^i\}_i\{z^j\vec{a}^j\vec{x}^j\}_j} \prod_i \tilde{p}(\vec{a}_i^i \vec{b}_i^i c^i | \vec{x}_i^i \vec{y}_i^i z^i \lambda_{A_i}) \times \prod_j \tilde{p}(\vec{a}_j^j \vec{b}_j^j c^j | \lambda_{B_j} | \vec{x}_j^j \vec{y}_j^j z^j) \\
&= \sum_{\substack{\{c^i\vec{b}^i\vec{y}^i\}_i\{\vec{c}^j\vec{b}^j\vec{y}^j\}_j \\ \{z^i\vec{a}^i\vec{x}^i\}_i\{z^j\vec{a}^j\vec{x}^j\}_j}} \alpha_{\{c^i\vec{b}^i\vec{y}^i\}_i\{\vec{c}^j\vec{b}^j\vec{y}^j\}_j}^{\{z^i\vec{a}^i\vec{x}^i\}_i\{z^j\vec{a}^j\vec{x}^j\}_j} \prod_i \text{Tr} \left[F_{\lambda_{A_i}}^{A_i} \tilde{M}_{\vec{a}_i^i \vec{b}_i^i c^i | \vec{x}_i^i \vec{y}_i^i z^i}^{A_i} \right] \times \prod_j \text{Tr} \left[H_{\lambda_{B_j}}^{B_j} \tilde{\rho}_{\vec{a}_j^j \vec{b}_j^j c^j | \vec{x}_j^j \vec{y}_j^j z^j}^{B_j} \right], \quad (\text{J57})
\end{aligned}$$

where the equality follows from Eqs. (J49) and (J50).

Plugging the operational identity in Eq. (J56) back into Eq. (J57), we thus have

$$p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) = \text{Tr} \left[(\otimes_j H_{\lambda_{B_j}}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} (\otimes_i F_{\lambda_{A_i}}^{A_i}) \right]. \quad (\text{J58})$$

Now, substituting Eqs. (J47), (J46), and (J58) into Eq. (J44a), and using the fact that we are considering every possible local measurement and preparation that are tomographically complete, we have

$$\begin{aligned}
\text{Tr} \left[\otimes_{j \in [m]} (M_{b_j|y_j}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\otimes_{i \in [n]} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right) \right] &= \sum_{\lambda_{\vec{B}} \lambda_{\vec{A}}} \text{Tr} \left[(\otimes_j H_{\lambda_{B_j}}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} (\otimes_i F_{\lambda_{A_i}}^{A_i}) \right] \prod_j \text{Tr} \left[M_{b_j|y_j}^{B_j} \sigma_{\lambda_{B_j}}^{B_j} \right] \prod_i \text{Tr} \left[G_{\lambda_{A_i}}^{A_i} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right] \\
\Rightarrow \mathcal{F}_{c|z}^{\vec{B}|\vec{A}}(\cdot) &= \sum_{\lambda_{\vec{A}} \lambda_{\vec{B}}} \otimes_j \sigma_{\lambda_{B_j}}^{B_j} \text{Tr} \left[(\otimes_j H_{\lambda_{B_j}}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} (\otimes_i F_{\lambda_{A_i}}^{A_i}) \right] \text{Tr} \left[(\otimes_i G_{\lambda_{A_i}}^{A_i})(\cdot) \right]. \quad (\text{J59})
\end{aligned}$$

Now let us prove the if part. If there exists a frame decomposition of the multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ then for any element of $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$, any element of a multi-measurement $M_{b|y}^B$ and any element of a multi-source $p(a|x)\rho_{a|x}^A$, one can define the individual ontological representations induced by the frame operators

$$p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) = \text{Tr} \left[(\otimes_j H_{\lambda_{B_j}}^{B_j}) \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} (\otimes_i F_{\lambda_{A_i}}^{A_i}) \right], \quad (\text{J60a})$$

$$p(a_i\lambda_{A_i}|x_i) := \text{Tr} \left[G_{\lambda_{A_i}}^{A_i} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \right], \quad (\text{J60b})$$

$$p(b_j|y_j\lambda_{B_j}) := \text{Tr} \left[\sigma_{\lambda_{B_j}}^{B_j} M_{b_j|y_j}^{B_j} \right], \quad (\text{J60c})$$

where the non-negativity and normalization are ensured by Eq. (J43), $\sigma_{\lambda_{B_j}}^{B_j}$ being a density operator and $\{G_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ being a POVM. Moreover, each is a linear functional on its corresponding quantum process.

One can also show that the effective ontological processes obtained by sequential compositions are also linear to the corresponding effective quantum processes obtained by sequential compositions. This is because $\{F_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$ and $\{G_{\lambda_{A_i}}^{A_i}\}_{\lambda_{A_i}}$, $\{\sigma_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$, and $\{H_{\lambda_{B_j}}^{B_j}\}_{\lambda_{B_j}}$ are a frame and its dual frame on V_{A_i} and V_{B_j} , respectively. In particular, for any choice of $\mathcal{I} \subseteq [1, \dots, n]$ and $\mathcal{J} \subseteq [1, \dots, m]$, one can verify by plugging in Eq. (J60a) to (J60c) and using the frame-dual-frame properties that

$$\begin{aligned}
p(c\vec{a}_{\mathcal{I}} \vec{b}_{\mathcal{J}} \lambda_{\vec{B}_{\mathcal{J}}} | z\vec{x}_{\mathcal{I}} \vec{y}_{\mathcal{J}} \lambda_{\vec{A}_{\mathcal{I}}}) &= \sum_{\lambda_{\vec{B}} \lambda_{\vec{A}}} p(c\lambda_{\vec{B}}|z\lambda_{\vec{A}}) \prod_{j \in \mathcal{J}} p(b_j|y_j\lambda_{B_j}) \prod_{i \in \mathcal{I}} p(a_i\lambda_{A_i}|x_i) \\
&= \text{Tr} \left[\otimes_{j \in \mathcal{J}} M_{b_j|y_j}^{B_j} \otimes_{j \in \mathcal{J}} H_{\lambda_{B_j}}^{B_j} \mathcal{F}_{c|z}^{\vec{B}|\vec{A}} \left(\otimes_{i \in \mathcal{I}} p(a_i|x_i) \rho_{a_i|x_i}^{A_i} \otimes_{i \notin \mathcal{I}} F_{\lambda_{A_i}}^{A_i} \right) \right]. \quad (\text{J60d})
\end{aligned}$$

Since all these ontological representations depend linearly on the corresponding physical processes and are non-negative, we have shown that the circuit is classically explainable. Since we allowed all possible factorizing

dual processes (i.e., all product multi-measurements and product multi-sources), this, in turn, implies that the multipartite multi-instrument $\{\{\mathcal{F}_{c|z}^{\vec{B}|\vec{A}}\}_c\}_z$ is classical by Definition 3'. ■

APPENDIX K: FRAME REPRESENTATION OF 2-COMBS

As we defined in the Introduction, a generic process, in the context of this paper, is simply a transformation with various input and output systems, which may be quantum or classical. Note that we did not specify the internal causal structure of a given quantum process. That is, here we did not stipulate any particular causal influence relations among the input and output systems [125]. For example, any multipartite quantum channel is treated as a monolithic object with multiple inputs and multiple outputs without constraints on its internal causal structure, so any input system may influence any output system. As a consequence, in our framework, a multipartite channel and an N -comb (as defined in Ref. [83]) are treated as the same type of process, even though they are physically distinct, with the latter subject to additional no-signaling constraints arising from its internal causal structure. Nevertheless, as we will see by considering frame representations of N -combs, although we never specify the internal causal structure of a given process, if an N -comb is classical, its corresponding ontological model nevertheless automatically satisfies the required no-signaling constraints.

Let us take a 2-comb $\mathcal{C}^{B_2 B_1 | A_2 A_1}$ as an example. For $\mathcal{C}^{B_2 B_1 | A_2 A_1}$ to be a valid 2-comb, it must be completely positive and trace preserving, and it must also satisfy a no-signaling constraint between input A_2 and output B_1 . That is, there exists a channel $\mathcal{E}^{B_1 | A_1}$ such that

$$\text{Tr}_{B_2} [\mathcal{C}^{B_2 B_1 | A_2 A_1}(\rho^{A_1} \otimes \sigma^{A_2})] = \mathcal{E}^{B_1 | A_1}(\rho^{A_1}) \quad (\text{K1})$$

for all states ρ^{A_1} and σ^{A_2} .

Now apply Theorem 6 to this 2-comb by viewing it as a structureless process. A classical 2-comb then must admit a frame representation of the form

$$\begin{aligned} \mathcal{C}^{B_2 B_1 | A_2 A_1}(\cdot) &= \sum_{\lambda_{A_1} \lambda_{A_2}} p(\lambda_{B_1} \lambda_{B_2} | \lambda_{A_1} \lambda_{A_2}) \\ &\quad \times \text{tr} \left[(G_{\lambda_{A_1}}^{A_1} \otimes G_{\lambda_{A_2}}^{A_2})(\cdot) \right] \sigma_{\lambda_{B_1}}^{B_1} \otimes \sigma_{\lambda_{B_2}}^{B_2}, \end{aligned} \quad (\text{K2})$$

where

$$\begin{aligned} p(\lambda_{B_1} \lambda_{B_2} | \lambda_{A_1} \lambda_{A_2}) \\ := \text{tr} \left[(H_{\lambda_{B_1}}^{B_1} \otimes H_{\lambda_{B_2}}^{B_2}) \mathcal{C}^{B_2 B_1 | A_2 A_1} (F_{\lambda_{A_1}}^{A_1} \otimes F_{\lambda_{A_2}}^{A_2}) \right] \geq 0 \end{aligned} \quad (\text{K3})$$

is its corresponding ontological representation.

One can now directly verify that $p(\lambda_{B_1} \lambda_{B_2} | \lambda_{A_1} \lambda_{A_2})$ satisfies the appropriate no-signaling constraint, i.e.,

$$\begin{aligned} &\sum_{\lambda_{B_2}} p(\lambda_{B_1} \lambda_{B_2} | \lambda_{A_1} \lambda_{A_2}) \\ &= \sum_{\lambda_{B_2}} \text{tr} \left[(H_{\lambda_{B_1}}^{B_1} \otimes H_{\lambda_{B_2}}^{B_2}) \mathcal{C}^{B_2 B_1 | A_2 A_1} (F_{\lambda_{A_1}}^{A_1} \otimes F_{\lambda_{A_2}}^{A_2}) \right] \\ &= \text{tr} \left[(H_{\lambda_{B_1}}^{B_1} \otimes \mathbb{1}^{B_2}) \mathcal{C}^{B_2 B_1 | A_2 A_1} (F_{\lambda_{A_1}}^{A_1} \otimes F_{\lambda_{A_2}}^{A_2}) \right] \\ &= \text{tr}_{B_1} \left[H_{\lambda_{B_1}}^{B_1} \mathcal{E}^{B_1 | A_1} (F_{\lambda_{A_1}}^{A_1}) \right] \\ &= p(\lambda_{B_1} | \lambda_{A_1}) \geq 0, \end{aligned} \quad (\text{K4})$$

where we used $\sum_{\lambda_{B_2}} H_{\lambda_{B_2}}^{B_2} = \mathbb{1}^{B_2}$ in the second equality and the no-signaling condition of the 2-comb in the third equality. The non-negativity of the final expression is automatic.

Therefore, although our framework does not assume any internal causal structure of a given process, and treats it only as a structureless object by specifying only its input and output systems, our framework can naturally be applied to those higher-order processes with fixed internal causal structure. Intuitively, this is because the corresponding no-signaling constraints for any process are automatically satisfied in its ontological representation due to the structure theorems we provide.

However, we do not claim that our framework can be directly generalized to treat arbitrary higher-order quantum processes. In particular, it is not clear how to define an ontological model for a process with indefinite causal order. Explicitly specifying the internal causal structure as an operational constraint might be important for studying arbitrary higher-order processes. We view this as an interesting direction for future research, which will likely require further investigation.

APPENDIX L: PROOF OF LEMMA 1

We now repeat and then prove Lemma 1. The connection between the prepare-transform-measure circuit for a channel and the corresponding bipartite prepare-measure circuit for the channel's Choi state is depicted in Fig. 28.

Lemma 1. The Choi state isomorphic to a given process is classical if and only if the process is classical.

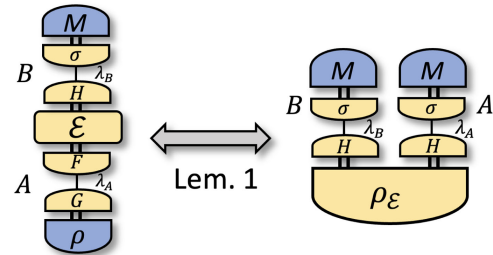


FIG. 28. Connecting a prepare-transform-measure circuit with a prepare-measure circuit on a bipartite system involving local measurements.

Proof. Let us first show that if a quantum channel $\mathcal{E}^{B|A}$ is classical, then so is its Choi state $\rho_{\mathcal{E}}^{BA} = (\mathcal{E}^{B|A} \otimes \mathbb{1}^A)|\Psi^+\rangle\langle\Psi^+|^{AA}$.

By Theorem 5, a classical channel admits a frame decomposition

$$\mathcal{E}^{B|A}(\cdot) = \sum_{\lambda_A \lambda_B} \sigma_{\lambda_B}^B \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A(\cdot)], \quad (\text{L1})$$

where $\{\sigma_{\lambda_B}^B\}_{\lambda_B}$ is a set of quantum states defining a frame on $V_B := \text{Span}[\{\mathcal{E}^{B|A}(\rho^A)\}]$, and $\{G_{\lambda_A}^A\}_{\lambda_A}$ is a POVM defining a frame on $V_A := \text{Span}[\{[\mathcal{E}^{B|A}]^\dagger(M^B)\}]$, and where $\{H_{\lambda_B}^B\}_{\lambda_B}$ and $\{F_{\lambda_A}^A\}_{\lambda_A}$ are sets of Hermitian operators forming the dual frames that satisfy $\sum_{\lambda_B} H_{\lambda_B}^B = \mathcal{P}_{V_B}(\mathbb{1}^B)$, $\text{Tr}[F_{\lambda_A}^A] = 1$ and moreover,

$$\text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \geq 0, \quad \forall \lambda_A, \lambda_B. \quad (\text{L2})$$

We can then write the Choi state $\rho_{\mathcal{E}}^{BA}$ of the quantum channel $\mathcal{E}^{A \rightarrow B}$ in terms of this frame decomposition as

$$\begin{aligned} \rho_{\mathcal{E}}^{BA} &= (\mathcal{E} \otimes \mathbb{1}^A)|\Psi^+\rangle\langle\Psi^+|^{AA} \\ &= \sum_{\lambda_A \lambda_B} \sigma_{\lambda_B}^B \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \text{Tr}[G_{\lambda_A}^A \otimes \mathbb{1}^A |\Psi^+\rangle\langle\Psi^+|^{AA}] \\ &= \frac{1}{d} \sum_{\lambda_A \lambda_B} \sigma_{\lambda_B}^B \otimes [G_{\lambda_A}^A]^T \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \\ &= \sum_{\lambda_A \lambda_B} \sigma_{\lambda_B}^B \otimes [G_{\lambda_A}^A]^T \text{Tr}[H_{\lambda_B}^B \otimes [F_{\lambda_A}^A]^T \rho_{\mathcal{E}}^{BA}], \end{aligned} \quad (\text{L3})$$

where all transposes are with respect to the basis defining $|\Psi^+\rangle := (1/\sqrt{d}) \sum_i |ii\rangle$, and we use the standard Choi identities $\text{Tr}_A[X \otimes \mathbb{1} |\Psi^+\rangle\langle\Psi^+|] = (1/d) X^T$ and $(1/d) \text{Tr}[Y^B \mathcal{E}^{B|A}(X^A)] = \text{Tr}[(Y^B \otimes [X^A]^T) \rho_{\mathcal{E}}^{BA}]$.

We can now rescale $\{G_{\lambda_A}^A\}_{\lambda_A}$ and $\{F_{\lambda_A}^A\}_{\lambda_A}$ to define a new frame and its dual frame

$$\sigma_{\lambda_A}^A := \frac{[G_{\lambda_A}^A]^T}{\text{Tr}[G_{\lambda_A}^A]}, \quad H_{\lambda_A}^A := [F_{\lambda_A}^A]^T \text{Tr}[G_{\lambda_A}^A], \quad (\text{L4})$$

such that $\{\sigma_{\lambda_A}^A\}_{\lambda_A}$ is a set of quantum states, and $\{H_{\lambda_A}^A\}_{\lambda_A}$ is a set of Hermitian operators such that $\sum_{\lambda_A} H_{\lambda_A}^A = \mathcal{P}_{V_A}(\mathbb{1}^A) = \mathbb{1}^A$ since $V_A = \text{Span}(\{[\mathcal{E}^{B|A}]^\dagger(M^B)\})$ must contain the identity operator $\mathbb{1}^A$ as the map $[\mathcal{E}^{B|A}]^\dagger$ is unital; moreover, the non-negativity conditions are respected since

$$\text{Tr}[H_{\lambda_B}^B \otimes H_{\lambda_A}^A \rho_{\mathcal{E}}^{BA}] = \frac{1}{d} \text{Tr}[G_{\lambda_A}^A] \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] \geq 0. \quad (\text{L5})$$

Therefore, by Theorem 4, $\rho_{\mathcal{E}}^{BA}$ is classical.

Now we prove the reverse direction. A classical bipartite state $\rho_{\mathcal{E}}^{BA}$ admits a frame decomposition

$$\rho_{\mathcal{E}}^{BA} = \sum_{\lambda_A \lambda_B} \text{Tr}[(H_{\lambda_B}^B \otimes H_{\lambda_A}^A) \rho_{\mathcal{E}}^{BA}] \sigma_{\lambda_B}^B \otimes \sigma_{\lambda_A}^A, \quad (\text{L6})$$

where $\{\sigma_{\lambda_B}^B\}$ and $\{\sigma_{\lambda_A}^A\}$ are sets of quantum states defining two respective frames on $V_{B(A)} = \text{Span}(\{\sigma_{\lambda_{B(A)}}^{B(A)}\})$, while $\{H_{\lambda_B}^B\}$, $\{H_{\lambda_A}^A\}$ are sets of Hermitian operators defining the respective dual frames and satisfying $\sum_{\lambda_B} H_{\lambda_B}^B = \mathcal{P}_{V_B}(\mathbb{1}^B)$, and $\sum_{\lambda_A} H_{\lambda_A}^A = \mathcal{P}_{V_A}(\mathbb{1}^A) = \mathbb{1}^A$ where

$$\text{Tr}[(H_{\lambda_B}^B \otimes H_{\lambda_A}^A) \rho_{\mathcal{E}}^{BA}] \geq 0, \quad \forall \lambda_B, \lambda_A. \quad (\text{L7})$$

Since a Choi state must satisfy $\text{Tr}_B[\rho_{\mathcal{E}}^{BA}] = (1/d) \mathbb{1}^A$, we have

$$\begin{aligned} \text{Tr}_B[\rho_{\mathcal{E}}^{BA}] &= \sum_{\lambda_A} \text{Tr}[(\mathcal{P}_{V_B}(\mathbb{1}^B) \otimes H_{\lambda_A}^A) \rho_{\mathcal{E}}^{BA}] \sigma_{\lambda_A}^A \\ &= \sum_{\lambda_A} \text{Tr}_A[H_{\lambda_A}^A \text{Tr}_B[\rho_{\mathcal{E}}^{BA}]] \sigma_{\lambda_A}^A \\ &= \sum_{\lambda_A} \frac{1}{d} \text{Tr}_A[H_{\lambda_A}^A] \sigma_{\lambda_A}^A = \frac{1}{d} \mathbb{1}^A, \end{aligned} \quad (\text{L8})$$

where the second equality is implied by the fact that $\mathcal{P}_{V_B}(\mathbb{1}^B)$ acts as an identity operator on the span of V_B , and the last equality holds since $\{H_{\lambda_A}^A\}_{\lambda_A}$ and $\{\sigma_{\lambda_A}^A\}_{\lambda_A}$ define a pair of frame and dual frame on V_A , which contains $\mathbb{1}^A$.

Moreover, the constraints in Eq. (L7) imply

$$\sum_{\lambda_B} \text{Tr}[(H_{\lambda_B}^B \otimes H_{\lambda_A}^A) \rho_{\mathcal{E}}^{BA}] = \frac{1}{d} \text{Tr}[H_{\lambda_A}^A] \geq 0, \quad \forall \lambda_A. \quad (\text{L9})$$

We can now reconstruct the quantum channel $\mathcal{E}^{B|A}$ from its Choi state $\rho_{\mathcal{E}}^{BA}$ as

$$\begin{aligned} \mathcal{E}^{B|A}(\cdot) &= d \text{Tr}_A[(\mathbb{1}^B \otimes (\cdot)^T) \rho_{\mathcal{E}}^{BA}] \\ &= d \sum_{\lambda_B \lambda_A} \text{Tr}[(H_{\lambda_B}^B \otimes H_{\lambda_A}^A) \rho_{\mathcal{E}}^{BA}] \text{Tr}[\sigma_{\lambda_A}^A (\cdot)^T] \sigma_{\lambda_B}^B \\ &= \sum_{\lambda_B \lambda_A} \text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(H_{\lambda_A}^A)^T] \text{Tr}[(\sigma_{\lambda_A}^A)^T (\cdot)] \sigma_{\lambda_B}^B. \end{aligned} \quad (\text{L10})$$

We then rescale $\{G_{\lambda_A}^A\}_{\lambda_A}$ and $\{F_{\lambda_A}^A\}_{\lambda_A}$ to define a new frame and its dual frame by

$$F_{\lambda_A}^A := \frac{(H_{\lambda_A}^A)^T}{\text{Tr}[H_{\lambda_A}^A]}, \quad G_{\lambda_A}^A := \text{Tr}[H_{\lambda_A}^A] (\sigma_{\lambda_A}^A)^T, \quad (\text{L11})$$

such that $\text{Tr}[F_{\lambda_A}^A] = 1$ and $\sum_{\lambda_A} G_{\lambda_A}^A = \sum_{\lambda_A} \text{Tr}[H_{\lambda_A}^A] (\sigma_{\lambda_A}^A)^T = \mathcal{P}_{V_A}(\mathbb{1}^A) = \mathbb{1}^A$ using Eq. (L8). Thus, we have

$$\text{Tr}[H_{\lambda_B}^B \mathcal{E}^{B|A}(F_{\lambda_A}^A)] = \frac{d}{\text{Tr}[H_{\lambda_A}^A]} \text{Tr}[(H_{\lambda_B}^B \otimes H_{\lambda_A}^A) \rho_{\mathcal{E}}^{BA}] \geq 0,$$

where $\{d/(\text{Tr}[H_{\lambda_A}^A])\} \geq 0$ is ensured by Eq. (L9).

Therefore, a quantum channel $\mathcal{E}^{B|A}$ with a classical Choi state $\rho_{\mathcal{E}}^{BA}$ must be classical. ■

Remark. To show that for any general process with several quantum inputs and/or outputs, classicality is preserved under the Choi isomorphism, it suffices to apply the frame and dual-frame rescaling rules in Eq. (L4) or Eq. (L11) to each quantum input or output. All the non-negativity and normalization constraints are preserved, and no additional modifications are required.

APPENDIX M: PROOF OF LEMMA 3

We now prove Lemma 3. The connection between the prepare-measure circuit for a source and the corresponding bipartite prepare-measure circuit for a bipartite state is depicted in Fig. 29.

Lemma 3. Whether or not a process is nonclassical is independent of whether one reconceptualizes its classical outputs as dephased quantum outputs.

Proof. We prove the claim for a source $\{p(c)\rho_c^A\}$, which can also be viewed as a bipartite state

$$\rho^{CA} := \sum_c p(c) |c\rangle\langle c|^C \otimes \rho_c^A \quad (\text{M1})$$

when the classical outcome is reconceptualized as the dephased quantum output C . The generalization to arbitrary quantum processes is straightforward.

For one direction, if the bipartite state ρ^{CA} is classical, following Proposition 2, then any source on A obtainable by a measurement on C is classical. In particular, measuring $\{M_c = |c\rangle\langle c|^C\}_c$ on ρ^{CA} yields the source $\{p(c)\rho_c^A\}$, which must therefore be classical.

For the other direction, if the source $\{p(c)\rho_c^A\}$ is classical, according to Theorem 2, there exists a frame decomposition such that

$$\rho_c^A = \sum_{\lambda} \text{Tr}[H_{\lambda}^A \rho_c^A] \sigma_{\lambda}^A, \quad (\text{M2})$$

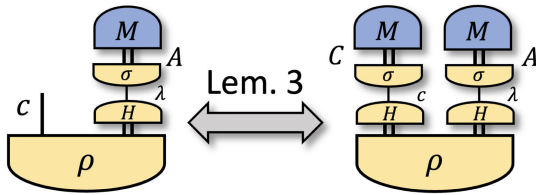


FIG. 29. Connecting a prepare-measure circuit for a source with a prepare-measure circuit on a bipartite system involving local measurements.

where $\{\sigma_{\lambda}^A\}_{\lambda}$ is a set of quantum states, and $\{H_{\lambda}^A\}$ is a set of Hermitian operators with $\sum_{\lambda} H_{\lambda}^A = \mathcal{P}_V(\mathbb{1}^A)$ with $V = \text{span}(\{\rho_c\}_c)$, and $\text{Tr}[H_{\lambda}^A \rho_c^A] \geq 0$ for all λ, c .

One can then define another pair of frame and dual frame as $\{\sigma_c^C := |c\rangle\langle c|^C\}_c$ and $\{H_c^C := |c\rangle\langle c|^C\}_c$ with $\text{Tr}[H_c^C \sigma_{c'}^C] = |\langle c|c'\rangle|^2 = \delta_{c,c'}$. With that, one can find a frame decomposition of ρ^{CA} as

$$\rho^{CA} = \sum_{c,\lambda} \text{Tr}[(H_c^C \otimes H_{\lambda}^A) \rho^{CA}] \sigma_c^C \otimes \sigma_{\lambda}^A. \quad (\text{M3})$$

Moreover, since $\text{Tr}[H_{\lambda}^A \rho_c^A] \geq 0$ for all λ , we have

$$\text{Tr}[(H_c^C \otimes H_{\lambda}^A) \rho^{CA}] = p(c) \text{Tr}[H_{\lambda}^A \rho_c^A] \geq 0, \quad \forall c, \lambda.$$

Therefore, the bipartite state ρ^{CA} is classical. ■

Remark. For any general process with several quantum inputs and outputs, classicality remains preserved when classical outcomes are reconceptualized as dephased quantum outputs. This follows from choosing the eigenbasis $\{|c\rangle\langle c|\}_c$ of each dephased system to serve simultaneously as the frame and its dual, ensuring that all associated non-negativity and normalization constraints are automatically respected.

-
- [1] An exception to this statement is that a unified proposal has been made for studying nonclassicality of common causes in multipartite scenarios [2–4].
 - [2] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, *Quantum* **4**, 280 (2020).
 - [3] D. Schmid, D. Rosset, and F. Buscemi, *Quantum* **4**, 262 (2020).
 - [4] D. Schmid, T. C. Fraser, R. Kunjwal, A. B. Sainz, E. Wolfe, and R. W. Spekkens, *Quantum* **7**, 1194 (2023).
 - [5] R. W. Spekkens, *Evidence for the epistemic view of quantum states: A toy theory*, *Phys. Rev. A* **75**, 032110 (2007).
 - [6] R. W. Spekkens, *Quasi-quantization: Classical statistical theories with an epistemic restriction*, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer Netherlands, Dordrecht, 2016) pp. 83–135.
 - [7] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, *Why interference phenomena do not capture the essence of quantum theory*, *Quantum* **7**, 1119 (2023).
 - [8] R. W. Spekkens, *Contextuality for preparations, transformations, and unsharp measurements*, *Phys. Rev. A* **71**, 052108 (2005).
 - [9] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, *A structure theorem for generalized-noncontextual ontological models*, *Quantum* **8**, 1283 (2024).
 - [10] D. Schmid, J. H. Selby, and R. W. Spekkens, *Unscrambling the omelette of causation and inference: The framework of causal-inferential theories*, *arXiv:2009.03297*.
 - [11] R. W. Spekkens, *The ontological identity of empirical indiscernibles: Leibniz's methodological principle and its significance in the work of Einstein*, *arXiv:1909.04628*.

- [12] R. W. Spekkens, *Negativity and contextuality are equivalent notions of nonclassicality*, *Phys. Rev. Lett.* **101**, 020401 (2008).
- [13] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens, *Characterization of noncontextuality in the framework of generalized probabilistic theories*, *PRX Quantum* **2**, 010331 (2021).
- [14] L. Hardy, *Quantum theory from five reasonable axioms*, [arXiv:quant-ph/0101012](https://arxiv.org/abs/quant-ph/0101012).
- [15] J. Barrett, *Information processing in generalized probabilistic theories*, *Phys. Rev. A* **75**, 032304 (2007).
- [16] M. D. Mazurek, M. F. Pusey, R. Kunjwal, K. J. Resch, and R. W. Spekkens, *An experimental test of noncontextuality without unphysical idealizations*, *Nat. Commun.* **7**, 11780 (2016).
- [17] S. Kochen and E. Specker, *The problem of hidden variables in quantum mechanics*, *Indiana Univ. Math. J.* **17**, 59 (1968).
- [18] R. W. Spekkens, *Contextuality for preparations, transformations, and unsharp measurements*, *Phys. Rev. A* **71**, 052108 (2005).
- [19] R. Kunjwal and R. W. Spekkens, *From statistical proofs of the Kochen-Specker theorem to noise-robust noncontextuality inequalities*, *Phys. Rev. A* **97**, 052110 (2018).
- [20] R. Kunjwal and R. W. Spekkens, *From the Kochen-Specker theorem to noncontextuality inequalities without assuming determinism*, *Phys. Rev. Lett.* **115**, 110403 (2015).
- [21] J. S. Bell, *On the Einstein Podolsky Rosen paradox*, *Physics* **1**, 195 (1964).
- [22] D. Schmid, H. Du, J. H. Selby, and M. F. Pusey, *Uniqueness of noncontextual models for stabilizer subtheories*, *Phys. Rev. Lett.* **129**, 120403 (2022).
- [23] F. Shahandeh, *Quantum computational advantage implies contextuality*, [arXiv:2112.00024](https://arxiv.org/abs/2112.00024).
- [24] D. Schmid and R. W. Spekkens, *Contextual advantage for state discrimination*, *Phys. Rev. X* **8**, 011015 (2018).
- [25] K. Flatt, H. Lee, C. R. I. Carceller, J. B. Brask, and J. Bae, *Contextual advantages and certification for maximum-confidence discrimination*, *PRX Quantum* **3**, 030337 (2022).
- [26] S. Mukherjee, S. Naonit, and A. K. Pan, *Discriminating three mirror-symmetric states with a restricted contextual advantage*, *Phys. Rev. A* **106**, 012216 (2022).
- [27] J. Shin, D. Ha, and Y. Kwon, *Quantum contextual advantage depending on nonzero prior probabilities in state discrimination of mixed qubit states*, *Entropy* **23**, 1583 (2021).
- [28] L. Catani, M. Leifer, D. Schmid, and R. W. Spekkens, *Reply to “comment on ‘why interference phenomena do not capture the essence of quantum theory’”*, [arXiv:2207.11791](https://arxiv.org/abs/2207.11791).
- [29] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, *Aspects of the phenomenology of interference that are genuinely nonclassical*, *Phys. Rev. A* **108**, 022207 (2023).
- [30] T. Giordani, R. Wagner, C. Esposito, A. Camillini, F. Hoch, G. Carvacho, C. Pentangelo, F. Ceccarelli, S. Piacentini, A. Crespi *et al.*, *Experimental certification of contextuality, coherence, and dimension in a programmable universal photonic processor*, *Sci. Adv.* **9**, ead4249 (2023).
- [31] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, *Contextuality without incompatibility*, *Phys. Rev. Lett.* **130**, 230201 (2023).
- [32] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, *Accessible fragments of generalized probabilistic theories, cone equivalence, and applications to witnessing nonclassicality*, *Phys. Rev. A* **107**, 062203 (2023).
- [33] A. Tavakoli and R. Uola, *Measurement incompatibility and steering are necessary and sufficient for operational contextuality*, *Phys. Rev. Res.* **2**, 013011 (2020).
- [34] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, *What is nonclassical about uncertainty relations?*, *Phys. Rev. Lett.* **129**, 240401 (2022).
- [35] M. Lostaglio, *Certifying quantum signatures in thermodynamics and metrology via contextuality of quantum linear response*, *Phys. Rev. Lett.* **125**, 230603 (2020).
- [36] N. E. Comar, D. Cius, L. F. Santos, R. Wagner, and B. Amaral, *Contextuality in anomalous heat flow*, *PRX Quantum* **6**, 030359 (2025).
- [37] M. Lostaglio, *Quantum fluctuation theorems, contextuality, and work quasiprobabilities*, *Phys. Rev. Lett.* **120**, 040602 (2018).
- [38] M. F. Pusey, *Anomalous weak values are proofs of contextuality*, *Phys. Rev. Lett.* **113**, 200401 (2014).
- [39] R. Kunjwal, M. Lostaglio, and M. F. Pusey, *Anomalous weak values and contextuality: Robustness, tightness, and imaginary parts*, *Phys. Rev. A* **100**, 042116 (2019).
- [40] V. P. Rossi, D. Schmid, J. H. Selby, and A. B. Sainz, *Contextuality with vanishing coherence and maximal robustness to dephasing*, *Phys. Rev. A* **108**, 032213 (2023).
- [41] R. Wagner, A. Camillini, and E. F. Galvão, *Coherence and contextuality in a Mach-Zehnder interferometer*, *Quantum* **8**, 1240 (2024).
- [42] R. Wagner, R. S. Barbosa, and E. F. Galvão, *Inequalities witnessing coherence, nonlocality, and contextuality*, *Phys. Rev. A* **109**, 032220 (2024).
- [43] R. D. Baldijão, R. Wagner, C. Duarte, B. Amaral, and M. T. Cunha, *Emergence of noncontextuality under quantum Darwinism*, *PRX Quantum* **2**, 030351 (2021).
- [44] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, *Preparation contextuality powers parity-oblivious multiplexing*, *Phys. Rev. Lett.* **102**, 010401 (2009).
- [45] A. Chailloux, I. Kerenidis, S. Kundu, and J. Sikora, *Optimal bounds for parity-oblivious random access codes*, *New J. Phys.* **18**, 045003 (2016).
- [46] A. Ambainis, M. Banik, A. Chaturvedi, D. Kravchenko, and A. Rai, *Parity oblivious d-level random access codes and class of noncontextuality inequalities*, *Quantum Inf. Process.* **18**, 111 (2019).
- [47] D. Saha, P. Horodecki, and M. Pawłowski, *State independent contextuality advances one-way communication*, *New J. Phys.* **21**, 093057 (2019).
- [48] S. A. Yadavalli and R. Kunjwal, *Contextuality in entanglement-assisted one-shot classical communication*, *Quantum* **6**, 839 (2022).
- [49] A. Hameedi, A. Tavakoli, B. Marques, and M. Bourennane, *Communication games reveal preparation contextuality*, *Phys. Rev. Lett.* **119**, 220402 (2017).

- [50] A. M. Fonseca, V. P. Rossi, R. D. Baldijão, J. H. Selby, and A. B. Sainz, *Robustness of contextuality under different types of noise as quantifiers for parity-oblivious multiplexing tasks*, *Phys. Rev. A* **111**, 022217 (2025).
- [51] M. Lostaglio and G. Senno, *Contextual advantage for state-dependent cloning*, *Quantum* **4**, 258 (2020).
- [52] P. Jokinen, M. Weilenmann, M. Plávala, J.-P. Pellonpää, J. Kiukas, and R. Uola, *No-broadcasting characterizes operational contextuality*, *Phys. Rev. Lett.* **133**, 240201 (2024).
- [53] M. S. Leifer and R. W. Spekkens, *Pre- and post-selection paradoxes and contextuality in quantum mechanics*, *Phys. Rev. Lett.* **95**, 200405 (2005).
- [54] C. Rochi Carceller, K. Flatt, H. Lee, J. Bae, and J. B. Brask, *Quantum vs noncontextual semi-device-independent randomness certification*, *Phys. Rev. Lett.* **129**, 050501 (2022).
- [55] M. S. Leifer and O. J. E. Maroney, *Maximally epistemic interpretations of the quantum state and contextuality*, *Phys. Rev. Lett.* **110**, 120401 (2013).
- [56] D. Schmid, *A review and reformulation of macroscopic realism: Resolving its deficiencies using the framework of generalized probabilistic theories*, *Quantum* **8**, 1217 (2024).
- [57] V. J. Wright and M. Farkas, *Invertible map between Bell nonlocal and contextuality scenarios*, *Phys. Rev. Lett.* **131**, 220202 (2023).
- [58] R. Kunjwal and R. W. Spekkens, *From the Kochen-Specker theorem to noncontextuality inequalities without assuming determinism*, *Phys. Rev. Lett.* **115**, 110403 (2015).
- [59] R. Kunjwal, *Contextuality beyond the Kochen-Specker theorem*, [arXiv:1612.07250](https://arxiv.org/abs/1612.07250).
- [60] R. Kunjwal, *Beyond the Cabello-Severini-Winter framework: Making sense of contextuality without sharpness of measurements*, *Quantum* **3**, 184 (2019).
- [61] R. Kunjwal, *Hypergraph framework for irreducible non-contextuality inequalities from logical proofs of the Kochen-Specker theorem*, *Quantum* **4**, 219 (2020).
- [62] R. Kunjwal, *Fine's theorem, noncontextuality, and correlations in Specker's scenario*, *Phys. Rev. A* **91**, 022108 (2015).
- [63] T. Gonda, R. Kunjwal, D. Schmid, E. Wolfe, and A. B. Sainz, *Almost quantum correlations are inconsistent with Specker's principle*, *Quantum* **2**, 87 (2018).
- [64] D. Schmid, R. W. Spekkens, and E. Wolfe, *All the non-contextuality inequalities for arbitrary prepare-and-measure experiments with respect to any fixed set of operational equivalences*, *Phys. Rev. A* **97**, 062103 (2018).
- [65] A. Chaturvedi, M. Farkas, and V. J. Wright, *Characterising and bounding the set of quantum behaviours in contextuality scenarios*, *Quantum* **5**, 484 (2021).
- [66] D. Schmid, R. D. Baldijão, J. H. Selby, A. B. Sainz, and R. W. Spekkens, *Noncontextuality inequalities for prepare-transform-measure scenarios*, [arXiv:2407.09624](https://arxiv.org/abs/2407.09624).
- [67] L. Catani, T. D. Galley, and T. Gonda, *Resource-theoretic hierarchy of contextuality for general probabilistic theories*, [arXiv:2406.00717](https://arxiv.org/abs/2406.00717).
- [68] M. F. Pusey, L. del Rio, and B. Meyer, *Contextuality without access to a tomographically complete set*, [arXiv:1904.08699](https://arxiv.org/abs/1904.08699).
- [69] D. Schmid, J. H. Selby, V. P. Rossi, R. D. Baldijão, and A. B. Sainz, *Shadows and subsystems of generalized probabilistic theories: When tomographic incompleteness is not a loophole for contextuality proofs*, *Quantum* **9**, 1880 (2025).
- [70] D. Schmid, J. H. Selby, and R. W. Spekkens, *Addressing some common objections to generalized noncontextuality*, *Phys. Rev. A* **109**, 022228 (2024).
- [71] M. J. Grabowecky, Christopher A. J. Pollack, A. R. Cameron, R. W. Spekkens, and K. J. Resch, *Experimentally bounding deviations from quantum theory for a photonic three-level system using theory-agnostic tomography*, *Phys. Rev. A* **105**, 032204 (2022).
- [72] Y. Zhang, Y. Ying, and D. Schmid, *Quantifiers and witnesses for the nonclassicality of measurements and of states*, [arXiv:2504.02944](https://arxiv.org/abs/2504.02944).
- [73] Y. Zhang, J. Spodek, D. Schmid, C. Reid, L. J. Morrison, T. Jennewein, K. J. Resch, and R. W. Spekkens, *A new paradigm for entanglement certification using noncontextuality inequalities*, [arXiv:2507.01122](https://arxiv.org/abs/2507.01122).
- [74] This usage is analogous to how the term “rectangular” is sometimes used to mean nonsquare and other times taken in the more inclusive sense that includes squares as a special case.
- [75] The analog for the geometric example would be to use the term rectangular in the more inclusive sense (as the mathematicians do) and to use the terms “nonsquare” (or even “intrinsically rectangular”) to describe the property of failing to be square.
- [76] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
- [77] D. Schmid, K. Ried, and R. W. Spekkens, *Why initial system-environment correlations do not imply the failure of complete positivity: A causal perspective*, *Phys. Rev. A* **100**, 022112 (2019).
- [78] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Quantum circuit architecture*, *Phys. Rev. Lett.* **101**, 060401 (2008).
- [79] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Probabilistic theories with purification*, *Phys. Rev. A* **81**, 062348 (2010).
- [80] P.-A. Melliès, *Functorial boxes in string diagrams*, in *International Workshop on Computer Science Logic* (Springer, New York, 2006), pp. 1–30.
- [81] The constraints arising from this kind of operational identity must be included in order to ensure that the ontological model explains not only the way the processes are used in one's circuit, but also any possible way those processes could be used. Alternatively, one can view their inclusion as ensuring that the model explains the statistics arising in pairs of runs of the experiment. Note that running multiple copies of such an experiment is always necessary to gather statistics (see also Appendix B).
- [82] C. J. Wood and R. W. Spekkens, *The lesson of causal discovery algorithms for quantum correlations: Causal explanations of Bell-inequality violations require fine-tuning*, *New J. Phys.* **17**, 033002 (2015).

- [83] G. Chiribella, G. M. D'Ariano, and P. Perinotti, *Theoretical framework for quantum networks*, *Phys. Rev. A* **80**, 022339 (2009).
- [84] H. M. Wiseman, S. J. Jones, and A. C. Doherty, *Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox*, *Phys. Rev. Lett.* **98**, 140402 (2007).
- [85] Y. Zhang, J. Zhang, and E. Chitambar, *Cost of simulating entanglement in steering scenarios*, *Quantum* **9**, 1902 (2025).
- [86] M. F. Pusey, *Negativity and steering: A stronger Peres conjecture*, *Phys. Rev. A* **88**, 032313 (2013).
- [87] F. Buscemi, *All entangled quantum states are nonlocal*, *Phys. Rev. Lett.* **108**, 200401 (2012).
- [88] B. Zjawin, D. Schmid, M. J. Hoban, and A. B. Sainz, *Quantifying EPR: The resource theory of nonclassicality of common-cause assemblages*, *Quantum* **7**, 926 (2023).
- [89] B. Zjawin, D. Schmid, M. J. Hoban, and A. B. Sainz, *The resource theory of nonclassicality of channel assemblages*, *Quantum* **7**, 1134 (2023).
- [90] M. F. Pusey, *Robust preparation noncontextuality inequalities in the simplest scenario*, *Phys. Rev. A* **98**, 022112 (2018).
- [91] C. H. Bennett and G. Brassard, *Quantum cryptography: Public key distribution and coin tossing*, *Theor. Comput. Sci.* **560**, 7 (2014).
- [92] O. Gühne, E. Haapasalo, T. Kraft, J.-P. Pellonpää, and R. Uola, *Colloquium: Incompatible measurements in quantum information science*, *Rev. Mod. Phys.* **95**, 011003 (2023).
- [93] We thank Thomas Galley for discussions in which this point was first realized.
- [94] J. H. Selby, D. Schmid, E. Wolfe, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, *Contextuality without incompatibility*, *Phys. Rev. Lett.* **130**, 230201 (2023).
- [95] R. F. Werner, *Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model*, *Phys. Rev. A* **40**, 4277 (1989).
- [96] H. M. Wiseman, S. J. Jones, and A. C. Doherty, *Steering, entanglement, nonlocality, and the Einstein-Podolsky-Rosen paradox*, *Phys. Rev. Lett.* **98**, 140402 (2007).
- [97] Y. Zhang and E. Chitambar, *Exact steering bound for two-qubit Werner states*, *Phys. Rev. Lett.* **132**, 250201 (2024).
- [98] M. J. Renner, *Compatibility of generalized noisy qubit measurements*, *Phys. Rev. Lett.* **132**, 250202 (2024).
- [99] R. W. Spekkens, D. H. Buzacott, A. J. Keehn, B. Toner, and G. J. Pryde, *Preparation contextuality powers parity-oblivious multiplexing*, *Phys. Rev. Lett.* **102**, 010401 (2009).
- [100] K. Modi, A. Brodutch, H. Cable, T. Paterek, and V. Vedral, *The classical-quantum boundary for correlations: Discord and related measures*, *Rev. Mod. Phys.* **84**, 1655 (2012).
- [101] S. Jevtic, M. Pusey, D. Jennings, and T. Rudolph, *Quantum steering ellipsoids*, *Phys. Rev. Lett.* **113**, 020402 (2014).
- [102] P. Zanardi, D. A. Lidar, and S. Lloyd, *Quantum tensor product structures are observable induced*, *Phys. Rev. Lett.* **92**, 060402 (2004).
- [103] We have not developed our approach for processes with indefinite causal order, as it is not clear how to define an ontological model for these.
- [104] C. H. Bennett, D. P. DiVincenzo, T. Mor, P. W. Shor, J. A. Smolin, and B. M. Terhal, *Unextendible product bases and bound entanglement*, *Phys. Rev. Lett.* **82**, 5385 (1999).
- [105] M. Navascués, E. Wolfe, D. Rosset, and A. Pozas-Kerstjens, *Genuine network multipartite entanglement*, *Phys. Rev. Lett.* **125**, 240505 (2020).
- [106] M. Seevinck and G. Svetlichny, *Bell-type inequalities for partial separability in n -particle systems and quantum mechanical violations*, *Phys. Rev. Lett.* **89**, 060401 (2002).
- [107] G. Tóth and O. Gühne, *Entanglement detection in the stabilizer formalism*, *Phys. Rev. A* **72**, 022340 (2005).
- [108] R. Spekkens, *Motility of the internal-external cut as a foundational principle*, PIRSA:18040073 (2018).
- [109] Defining nonclassicality via steering is usually done for the particular class of multi-sources that arise in steering scenarios, namely, those wherein there is no causal influence of the setting variable z on the quantum output A , even though a generic multi-source allows such an influence. One might wonder whether this nuance could impact the point we are trying to make here. Specifically, might refining the class of bipartite multi-states in the same way, namely, to those wherein there is no causal influence of z on A , change the assessment of classicality of such a multi-state? Not if the assessment is made based on the notion that entanglement is necessary for nonclassicality regardless of how the bipartite state is prepared. This latter notion strikes us as being the conventional view, and so the nuance in question does not seem to impact our conclusion.
- [110] We note that the set of channels in question would also be considered nonclassical if one took the notion of channel incompatibility [111] to be a sufficient condition for nonclassicality.
- [111] T. Heinosaari, T. Miyadera, and D. Reitzner, *Strongly incompatible quantum devices*, *Found. Phys.* **44**, 34 (2014).
- [112] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, *Reconstruction of Gaussian quantum mechanics from Liouville mechanics with an epistemic restriction*, *Phys. Rev. A* **86**, 012103 (2012).
- [113] M. Plávala and O. Gühne, *Contextuality as a precondition for quantum entanglement*, *Phys. Rev. Lett.* **132**, 100201 (2024).
- [114] V. J. Wright and M. Farkas, *Invertible map between Bell nonlocal and contextuality scenarios*, *Phys. Rev. Lett.* **131**, 220202 (2023).
- [115] A. Tavakoli and R. Uola, *Measurement incompatibility and steering are necessary and sufficient for operational contextuality*, *Phys. Rev. Res.* **2**, 013011 (2020).
- [116] Y. Zhang, Y. Ying, D. Schmid, and R. W. Spekkens, *Towards resource theory of nonclassical processes: Compositional structure of nonclassical processes* (unpublished).
- [117] B. Coecke, T. Fritz, and R. W. Spekkens, *A mathematical theory of resources*, *Inf. Comput.* **250**, 59 (2016).
- [118] F. Riesz, *Démonstration nouvelle d'un théorème concernant les opérations fonctionnelles linéaires*, *Ann. Sci. École Norm. Supér.* **31**, 9 (1914).
- [119] C. Ferrie and J. Emerson, *Frame representations of quantum mechanics and the necessity of negativity in*

- quasi-probability representations*, *J. Phys. A* **41**, 352001 (2008).
- [120] M. D. Mazurek, M. F. Pusey, K. J. Resch, and R. W. Spekkens, *Experimentally bounding deviations from quantum theory in the landscape of generalized probabilistic theories*, *PRX Quantum* **2**, 020302 (2021).
- [121] E. Nielsen, J. K. Gamble, K. Rudinger, T. Scholten, K. Young, and R. Blume-Kohout, *Gate set tomography*, *Quantum* **5**, 557 (2021).
- [122] O. Christensen, *An Introduction to Frames and Riesz Bases*, Applied and Numerical Harmonic Analysis (Springer International Publishing, New York, 2016).
- [123] P. Busch, *Quantum states and generalized observables: A simple proof of Gleason's theorem*, *Phys. Rev. Lett.* **91**, 120403 (2003).
- [124] M. Horodecki, P. W. Shor, and M. B. Ruskai, *Entanglement breaking channels*, *Rev. Math. Phys.* **15**, 629 (2003).
- [125] For more discussions on the internal causal structure of a general circuit, see Ref. [126] Appendix F.1.
- [126] Y. Yīng, M. C. Alañón, D. Centeno, J. Surace, M. M. Ansanelli, R. Liu, D. Schmid, and R. W. Spekkens, *On whether quantum theory needs complex numbers: The foil theories perspective*, [arXiv:2506.08091](https://arxiv.org/abs/2506.08091).