PHYSICAL REVIEW A **109**, 042211 (2024)

Every nonsignaling channel is common-cause realizable

Paulo J. Cavalcanti,* John H. Selby,† and Ana Belén Sainz‡
International Centre for Theory of Quantum Technologies, University of Gdańsk, 80-309 Gdańsk, Poland

(Received 15 December 2023; accepted 8 March 2024; published 17 April 2024)

In this work we show that the set of nonsignaling resources of a locally tomographic generalized probabilistic theory (GPT), such as quantum and classical theory, coincides with its set of GPT-common-cause realizable resources, where the common causes come from an associated GPT. From a causal perspective, this result provides a reason for, in the study of resource theories of common-cause processes, taking the nonsignaling channels as the resources of the enveloping theory. This answers a critical open question in Schmidt *et al.* [Quantum **5**, 419 (2021)]. An immediate corollary of our result is that every nonsignaling assemblage is realizable in a GPT, answering in the affirmative the question posed in Cavalcanti *et al.* [npj Quantum Inf. **8**, 76 (2022)].

DOI: 10.1103/PhysRevA.109.042211

I. INTRODUCTION

There has been a great deal of recent interest in the study of resource theories [1] in which the free operations are either local operations and shared randomness (LOSR) [2–7], for the purposes of studying nonlocality and entanglement, or local operations and shared entanglement (LOSE) [8], for the purposes of studying postquantum nonlocality. In particular, it has been shown that these can be studied in a type-independent manner [8–10] such that resources of various types (entangled states, nonlocal boxes, steerable assemblages, etc.) can be treated in a uniform and unified way. These resource theories are motivated by the idea that the best way to understand Bell's theorem is from the perspective of causal models [4,11], and that the lesson to be learnt from Bell's theorem is that we need an intrinsically quantum notion of causality and of common causes [4,12].

Defining a resource theory requires a specification of both a *free* and an *enveloping* theory [1]. The free theory specifies the things that can be done effectively without cost, while the enveloping theory specifies the things that can be done irrespective of cost. While in the study of LOSE and LOSR it is clear how the free theory should be defined, it is not clear how the *enveloping theory* should be defined [8]. There are two options for this, each of which has pros and cons. On the one hand, we have the choice which is typically made, which is to use the enveloping theory which describes *nonsignaling* resources. The benefit of this choice is that it is mathematically simple to characterize, since in the cases of interest so far the set of such resources can often be expressed in a computationally easy way (polytope or semidefinite program) [13,14]. Its downside, however, is that this enveloping theory is not so well motivated from a causal perspective—it makes sense to say that resources should be nonsignaling, but why should *all* nonsignaling resources be considered? On the other hand, we can take the enveloping theory to describe arbitrary *common cause* resources, typically described using the framework of generalized probabilistic theories (GPTs) subsuming classical and quantum common causes as special cases. The benefit of this approach is that it is conceptually well motivated, from the causal perspective [4]. Its downside, however, is that providing a clean mathematical characterization of this enveloping theory is an open problem. The characterization and the relationship between these two options were cleanly articulated as an open question in Ref. [8, Open Question 1].

In this paper we resolve the tension between these two choices, by showing that these two options actually coincide. This means that we get the benefits of both approaches with none of the downsides. It is well established that every common-cause realizable resource is nonsignaling, so here we just focus on the converse direction. In particular, we show that there exists a GPT in which all nonsignaling resources of a target locally tomographic GPT, such as quantum theory, can be realized in a common-cause setting. On the one hand, we can view this result as providing a clear characterization of the set of GPT-realizable resources. On the other hand, we can also view it as providing a principled justification, backed by the causal perspective, for choosing the set of nonsignaling resources as the enveloping theory in resource theories of common-cause processes. We moreover show that this result holds not only in the bipartite case, which has so far dominated the literature, but also in the general multipartite scenario, thereby setting the stage for explorations of multipartite generalizations of LOSR and LOSE resource theories. A corollary of this result answers one of the open questions posed in Ref. [15]; namely, it shows that indeed any nonsignaling assemblage can be given a GPT-common-cause explanation.

The scheme by which we build the GPT where all nonsignaling resources can be realized in a common-cause setting differs from the standard approach to GPT

*paulojcvf@gmail.com
†john.h.selby@gmail.com
‡ana.sainz@ug.edu.pl

construction in the literature. Usually GPTs are constructed by making reference to the geometry of their states, effects, and transformations spaces, requiring, for example, that they are convex subsets of linear spaces (see, e.g., Ref. [16]). Here, instead of putting emphasis on the geometry, we focus our attention on compositionality, that is, we take a process-theoretic [17–20] approach to constructing one GPT from another. By focusing on the compositional properties of the theory, our method also has the potential to be applied to other problems.

To be more formal, let us define a *common-cause completion* of a given GPT G as an supertheory of G which can realize all of the nonsignaling resources G in a common-cause scenario. If some theory is the common-cause completion of itself, then we call it *common-cause complete*, in contrast to quantum and classical theory, which have nonsignaling resources which cannot be realized in common cause scenarios, being therefore common-cause incomplete. In this paper, we define a common-cause completion map, $\mathcal{C}$, which takes an arbitrary tomographically local GPT G as an input, and gives a common-cause completion of it, $\mathcal{C}$[G], as output. Specifically, this means that G is a full subtheory of $\mathcal{C}$[G] and that every nonsignaling resource in G can be realized with only common-cause resources in $\mathcal{C}$[G]. Proving the existence of such a common-cause completion map demonstrates the main claims of this paper: (1) all nonsignaling resources in G are common-cause realizable in the GPT $\mathcal{C}$[G], and so the nonsignaling resources in G coincide with its GPT-common-cause realizable processes and (2) every nonsignaling assemblage in G is realizable in an EPR scenario in $\mathcal{C}$[G].

II. GENERALIZED PROBABILISTIC THEORIES (GPTs)

In this section we provide a concise overview of generalized probabilistic theories (GPTs) [16,21], emphasizing their compositional attributes. We provide a brief introduction here, and refer readers to, for example, Refs. [22,23] for more details. Specifically, we are following the formalism of Refs. [19,20].

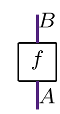
Conceptually, a GPT is a theory about experiments that assigns probabilities to observation events, equipped with a compositional structure that mirrors the possibility we have to perform actions sequentially or in parallel. Formally speaking, the compositional aspects of the theory are captured by the fact that a GPT is a (strict) symmetric monoidal category (SMC) (see Appendix A). The probabilistic aspects are captured by the fact that we have a classical (stochastic) interface with the full theory in order to represent outcomes and control variables; formally, this means that we have the SMC Stoch (Sec. II B) as a full subtheory. This leads to a convex structure (Sec. II C 2) on the sets of processes with a given input and output, and allows us to define a notion of tomography (Sec. II C 3). Finally, we capture the requirement that the theory interact well with relativistic causal structure, by demanding the existence of unique discarding maps (Sec. II C 4).

In the rest of this section, we will introduce the diagrammatic notation used throughout this work, and discuss the defining features of a GPT that we mentioned above.

A. Diagrammatic notation

An interesting feature of SMCs is that they have a diagrammatic representation with which we can perform every calculation that we could using their axiomatic definition [24–26]. In the context of GPTs, we can represent their processes as boxes with input and output wires, and encode the composition of these processes by how they are wired together.

In the diagrammatic notation, each wire is named to represent a system type, and we follow the convention where those connected to the bottom of the boxes represent the input types of the process, while those at the top are the outputs. Note that this means that, in our convention, "time" in the diagrams flows from the bottom up. In this way we can represent a process $f : A \to B$, that takes a system of type A to a system of type B, as follows:

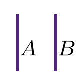
$f : A \to B \quad \doteq$  , (1)

where we are using $\doteq$ to indicate the translation from one notation into another.

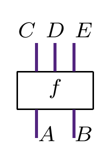
We often omit wire labels for simplicity and/or use colors to encode certain information about the system type. For instance, in this paper we will use

 (2)

where, for example, the first of these represents a classical system of unspecified dimension, and the meaning of the others will be explained in Sec. IV. To represent composite types such as $A \otimes B$, we just put their wires side by side, as in

$A \otimes B \quad \doteq$  . (3)

Using this notation for composite systems, a process with composite input or output wires is depicted as having multiple input and output wires, e.g.,

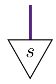
$f : A \otimes B \to C \otimes D \otimes E \quad \doteq$  . (4)

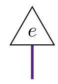
One system type that every GPT must contain is the trivial system, which corresponds to having no system at all. We refer to it in text as I. Since the trivial system is the unit for parallel composition (i.e., the monoidal unit of the symmetric monoidal category), we have $A \otimes I = A = I \otimes A$, diagrammatically, and I is represented by empty space:

$I \quad \doteq$. (5)

States and effects can be seen as preparation and observation procedures, respectively, which are processes that start and end in the trivial system, i.e., they must not have input or output wires respectively. For example, if s is a state and e

an effect, then we denote them as

$$s : I \to A \quad \doteq$$ 

$$\text{and} \quad e : A \to I \quad \doteq$$ 

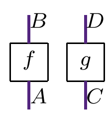
. (6)

There can also be processes with both input and output as the trivial system, $p : I \to I$, which are represented by diagrams without open wires. The compositional properties of the SMC imply that diagrams of this kind can be composed together with a multiplicative structure, and hence can be called numbers. For instance, we could have

$$1 : I \to I \quad \doteq$$ 

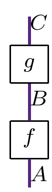
. (7)

Finally, we represent the parallel composition $\otimes$ of processes by drawing their boxes side by side, and their sequential composition $\circ$ by connecting the input and output wires of matching types. That is, for $f : A \to B$ and $g : C \to D$,

$$f \otimes g : A \otimes C \to B \otimes D \quad \doteq$$ 

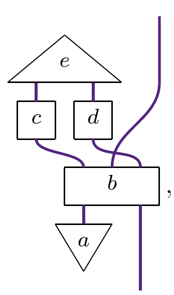
, (8)

and for $f : A \to B$ and $g : B \to C$,

$$g \circ f : A \to C \quad \doteq$$ 

. (9)

One example of a more complex diagram is



, (10)

where we omit the labels of the wires, but it should be understood that connections are allowed only when types match.

This notation and the rules for composing diagrams are common to all (strict) symmetric monoidal categories. Now it remains to discuss features that are shared only by those who can be considered as GPTs. Since one of the ingredients of a GPT is that they contain Stoch as a full subtheory, we start from the definition of that theory.

B. Example: Classical stochastic maps

As we mentioned, any GPT must have Stoch as a full subtheory. The simplest possible GPT then is the one that contains nothing else (if the other properties are satisfied, of course, which is the case).

In order to define Stoch, all we have to do is to define what concrete mathematical objects correspond to its system types, states, effects, transformations, and composition rules (parallel and sequential composition). We organize this information in Table I.

TABLE I. Elements in the definition of Stoch.

Element	Definition	Example
System types	Real vector spaces	$\mathbb{R}^2$
States	Probability column vectors	$\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$
Effects	Row vectors whose all entries are equal to 1	$\begin{pmatrix} 1 & 1 \end{pmatrix}$
Transformations	Stochastic matrices	$\begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}$
Sequential composition	Matrix multiplication	$\begin{pmatrix} 1/2 & 1/3 \\ 1/2 & 2/3 \end{pmatrix}\begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$
Parallel composition	Kronecker product (or tensor product)	$\begin{pmatrix} 1/2 \\ 2/3 \end{pmatrix} \otimes \begin{pmatrix} 1/2 \\ 1/2 \end{pmatrix}$

Because we require that GPTs have this theory as a full subtheory, it will act as an interface to provide the GPT with the probabilistic interpretation that we need. For example, in this framework we describe a measurement as a process from a general system to a system in Stoch.

C. Defining properties of causal GPTs

Not every SMC can be considered as being a hypothetical theory of physics. In this section we characterize those that can. In particular, what we are looking for with this characterization is to use Stoch as an interface to the theory that enables us to make statistical predictions in a manner coherent with its compositional structure, and where we can characterize the objects by the statistics that they can generate.

The additional features that an SMC has to satisfy in order to be a causal GPT are the following:

(1) The SMC contains Stoch as a *full subtheory*.

(2) There is a *convex structure* compatible with the one from Stoch.

(3) There is a notion of *tomography*.

(4) There is a *unique effect* associated with each system type.

In this paper we further focus on GPTs that satisfy the following additional property:

(5) The theory is *locally tomographic*.

We now discuss each of those points in turn.

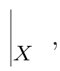
1. Stoch is a full subtheory

This means that all of the systems from Stoch and all of the processes from Stoch are also in the GPT, and, moreover, that when we compose these systems and processes in the GPT this matches the composition in Stoch [19]. Moreover, if we have a process in the GPT which has only inputs and outputs coming from Stoch, then this must be a process coming from Stoch.

The importance of that, is that inside a GPT, we can take the maps that go from a classical system (i.e., a system interpreted as a system of Stoch) to another one as a stochastic process. Then these processes, with all their internal probabilities, provide a probabilistic interpretation to the diagrams. Note that if it were not a *full* subtheory, then there would necessarily be situations in which the theory failed to make sensible probabilistic predictions, for example, giving negative probabilities for measurement outcomes.

For example, suppose we have a state of some general system in the GPT, A, then a (destructive) measurement for A would be a process with A as an input and some system X in Stoch as an output, when we compose these we are left with a process which must be a state in Stoch, namely, a probability distribution. It is precisely these probability distributions which encode the probabilistic predictions of the GPT.

We denote the systems coming from the subtheory Stoch as



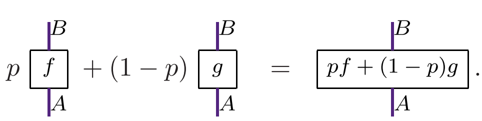
, (11)

where we use a thin gray wire to distinguish the systems in the subtheory from generic systems in the GPT.

Note that in many other approaches to GPTs, the probabilities are encoded as scalars in the theory. In the approach we take here this is not the case, as, in particular, we find here that there is a unique scalar, the number 1. Instead, we obtain probability distributions over measurement outcomes via the states of the subtheory, Stoch. For example, this is what we obtain when we compose a state of a generic system in sequence with a measurement on that system.

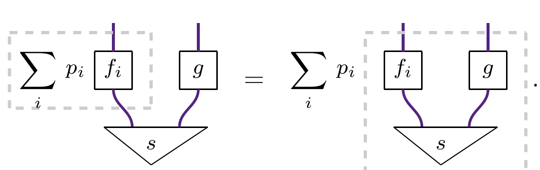
2. *Convex structure*

In order to naturally express statistical mixtures in the GPTs, we require them to be closed under convex mixtures of processes of matching input and output types. We require further that this composition is consistent with the convex composition from Stoch [19]. To start illustrating that, note that if we have $f : A \to B$ and $g : A \to B$, there must exist some $pf + (1-p)g : A \to B$ in the theory where we denote this as



. (12)

Note that these combinations are allowed only when the input-output systems are the same for each of the combined processes. Moreover, these must distribute over diagrams; i.e., they must satisfy, for example:



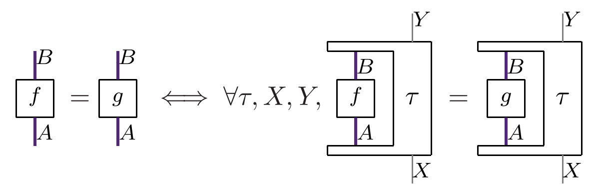
. (13)

Finally, these convex combinations must match up with the standard notion of convex-combinations when specialized to the subtheory Stoch. This ensures that we can consistently view these convex combinations as describing our classical uncertainty about which process is happening.

3. *Tomography*

The next requirement that a GPT must satisfy is to have a notion of tomography [27]. What that means is that we should be able to characterize its elements, i.e., the states, effects, and transformations, by the statistics that they are capable of generating. In this way, an experimentalist would be able to characterize the theoretical objects describing their experiment by connecting the statistics to the probabilities that the theory predicts.

To have a notion of tomography of processes, we need to always be able to establish equalities between them by looking at the statistics that they can generate. In a GPT, this means the following: we require that if it is the case that whenever we swap the process $f : A \to B$ by the process $g : A \to B$ in any diagram that represents a stochastic map, that map is kept unchanged, then it must be that $f = g$, i.e.,

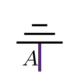


. (14)

Here we are using τ to represent an arbitrary diagram that, after inserting f in some specific spot thereof, has only classical inputs and outputs left, and so is a process in Stoch. Note that this includes the case where any of the input-output wires of τ are the trivial system, because the trivial system is a classical (that is, Stoch) system. This condition can be phrased in the following way: two processes f and g from A to B are equal [left-hand side of Eq. (14)] if and only if they are *operationally equivalent* [right-hand side of Eq. (14)].

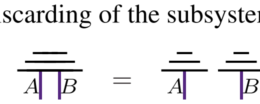
4. *Causality*

In this work we are interested in GPTs that are causal [28,29]. By that, we mean that for each system type A, there is a unique effect that we can think of as discarding, or simply ignoring, a given system. This property is called causality because it can be used to impose compatibility of the GPT with a relativistic causal structure [30]. When the theory satisfies causality, we use a special diagram to denote the unique (for each system A) discarding effect:



. (15)

Note that the uniqueness of the discarding effects is given for each fixed system type. In particular, this means that for composite systems the discarding is obtained by parallel composition of the discarding of the subsystems:

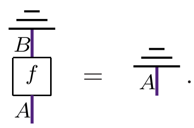


. (16)

The discarding effects will be used in the next section to define nonsignaling channels for a general GPT, just like the trace is in quantum theory.

The fact that there is a unique effect immediately means that all of the processes are discard-preserving [29].

Definition II.1 (Deterministic, or Discard-Preserving, Process). A process $f : A \to B$ is deterministic if it is discard preserving, that is,

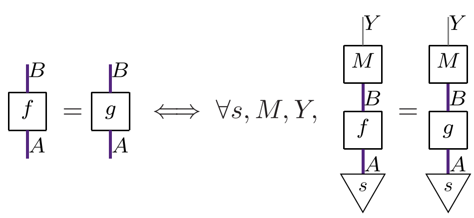


(17)

In quantum theory, since discarding is the trace operation, this corresponds to the trace-preserving property. That is, the formalism that we are using here is the analog of working with only CPTP maps rather than working with CPTNI maps. Typically, CPTNI maps are used to describe the potential outcomes of some measurement, and we can instead equally well work only with CPTP maps, by instead considering all possible outcomes at once, and keeping track of which outcome occurred by means of an auxiliary classical system.

5. Local tomography

In this work, we are interested in GPTs that satisfy a stricter notion of tomography. We require that the tomography of the processes can be done by evaluating the probabilities produced by local effects, that is, we require our GPT to satisfy local tomography [21]. This is expressed diagrammatically by the following:



(18)

where s is an arbitrary state of A, Y is an arbitrary classical system, and M is an arbitrary measurement of B. Note that this is taking a particular, less general, shape for τ in the definition of tomography.

Remark II.2. A very convenient fact about locally tomographic GPTs is that they are all subtheories of $\mathbb{R}$*Linear* [31–33] (Example in Appendix A), in the sense that all of the processes of the former are in the latter (or more rigorously, there is an injective map between their processes and system types), and they compose according to $\mathbb{R}$*Linear* compositional rules. This will come in handy, as in our construction we will use the fact that our GPT is one of $\mathbb{R}$*Linear*'s subtheories to write its processes in a mathematically concrete way. In particular, both classical and quantum theory satisfy local tomography, and therefore are also subtheories of $\mathbb{R}$*Linear*.

Now that we are done discussing the structure of the generalized probabilistic theories, we can proceed and focus on the properties of the processes that we are interested in investigating inside those theories. Namely, we can talk about the nonsignaling channels.

III. CHANNELS IN GENERALIZED PROBABILISTIC THEORIES

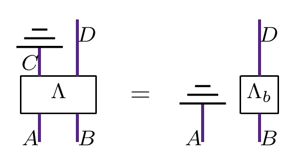
In this section we discuss, in the context of generalized probabilistic theories, the two classes of channels of interest for this paper: the nonsignaling channels, and the common-cause channels (which form a subset of the nonsignaling channels, as we will see).

A. Nonsignaling channels

A practical starting point to understand what nonsignaling channels in GPTs are is to remind ourselves of what they are in quantum or classical theory.

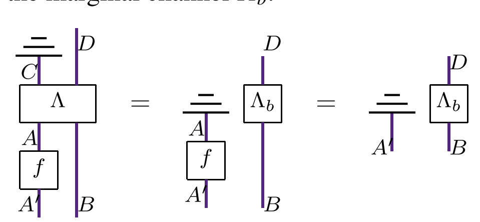
Quantum channels are formally completely positive trace-preserving maps on density matrices, and specify ways in which quantum systems can be transformed. The properties of quantum channels are widely studied in the literature [34], and of particular interest are the quantum channels that satisfy a form of the nonsignaling principle [35], introduced first by Beckman *et al.* [36] in bipartite setups. These nonsignaling quantum channels are sometimes referred to as "causal channels" [37] and do not permit superluminal quantum (nor classical) communication between two parties, i.e., two wings of the experiment. Nonsignaling channels were discussed in the context of multipartite setups by Schumacher and Westmoreland [37].

In general theories—not necessarily quantum or classical—one can also define the concept of a channel as a transformation in the theory that is discard-preserving (Definition II.1), that is, one that preserves, on any state, the result of the application of the discarding process. In this context, we can talk about the property of a channel being nonsignaling. In this section we present a convenient definition of nonsignaling channels in the diagrammatic language that we presented in Sec. II. Specifically, we want to diagrammatically represent the idea that no information can flow between the parties. Consider, for example, a bipartite process $\Lambda : A \otimes B \to C \otimes D$. If by discarding system C the resulting process $A \otimes B \to D$ is such that changing system A does not produce any changes in system D, then Λ cannot signal from the AC wing of the experiment to the BD wing of the experiment. In other words, we say that $\Lambda : A \otimes B \to C \otimes D$ is nonsignaling from AC to BD if and only if



(19)

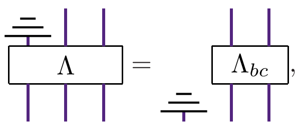
where $\Lambda_b : B \to D$ is a valid channel within the theory [29]. Note, in particular, that this implies that the application of any deterministic process (Definition II.1) in the AC wing does not change the marginal channel Λ_b:



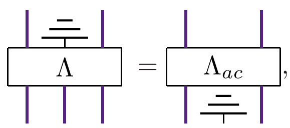
(20)

hence, no information can flow from the AC wing to the BD wing of the experiment. A channel is then said to be nonsignaling when it satisfies that property in both directions between the wings of the experiment.

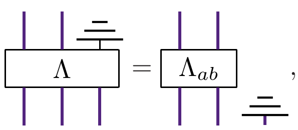
So far we have presented the case of bipartite nonsignaling channels, but the notion of a multipartite nonsignaling channel has also been defined in the literature [37]. Here we present a convenient diagrammatic definition of multipartite nonsignaling channels. In order to define the multipartite generalization of this condition we need a convenient way to represent discarding an arbitrary subset of the outputs. To see why, suppose that Λ is a tripartite channel. If we want to guarantee that no information can flow from any of the subsystems to any other, we need to have that



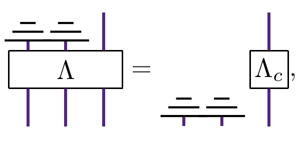
(21)



(22)



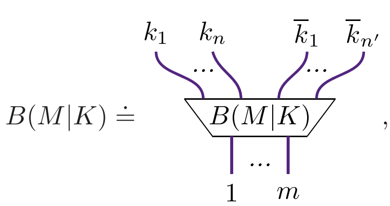
(23)



(24)

and so on. It is easy to see that this can become quite complex quickly as we increase the number of parties. In order to capture this in a succinct diagrammatic form, we need a notation which allows us to describe discarding an arbitrary subset of the outputs (or inputs), and for this purpose we first introduce a bipartitioning processes as follows.

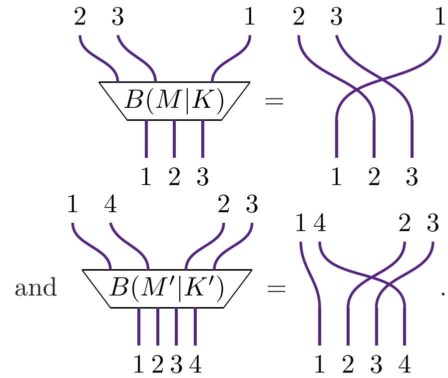
Definition III.1 (Bipartitioning processes $B(K)$). Given a set $M = \{1, \dots, m\}$ take a labeled subset $K = \{k_1, \dots, k_n\} \subset M$ and its complement $\overline{K} = \{\overline{k}_1, \dots, \overline{k}_{n'}\} = M \setminus K$, where $n + n' = m$. Then the bipartitioning process $B(K)$ is the permutation which takes $(1, \dots, m)$ to $(k_1, \dots, k_n, \overline{k}_1, \dots, \overline{k}_{n'})$. Diagrammatically we represent this by



(25)

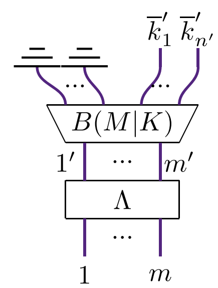
where we are using numbers, instead of system type names, to refer to the wires for the sake of clarity.

For example, if we take $M = \{1, 2, 3\}$ and $K = \{2, 3\}$, or $M' = \{1, 2, 3, 4\}$ and $K' = \{1, 4\}$ then we have, respectively,



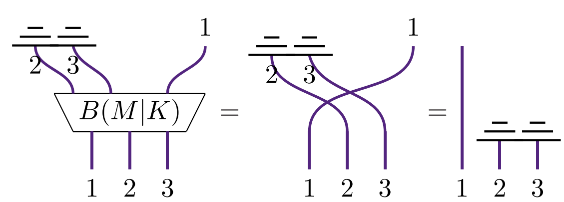
(26)

We can then use this bipartitioning operation to concisely notate discarding some subset K of the outputs M of a channel Λ,



(27)

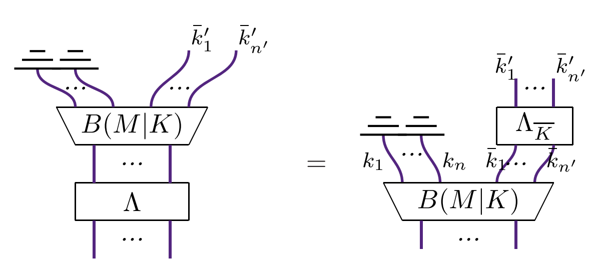
which in quantum theory would represent the partial trace $\mathrm{tr}_K(\Lambda)$, up to a permutation of the surviving systems. For example, in the tripartite case we can represent discarding the second and third outputs by



(28)

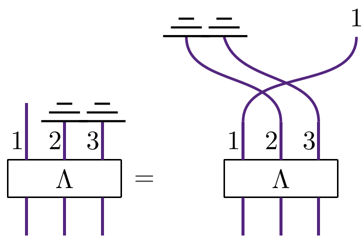
We can now present the definition of multipartite nonsignaling channels in a succinct diagrammatic form.

Definition III.2 (Nonsignaling channel). An m-partite channel $\Lambda : \otimes_{i=1}^{m} i \to \otimes_{i'=1}^{m} i'$ is nonsignaling iff for all labeled subsets $K \subset \{1, \dots, m\}$, there exists a channel $\Lambda_{\overline{K}} : \otimes_{i=1}^{n'} \overline{k}_i \to \otimes_{i=1}^{n'} \overline{k}'_i$, with $\overline{K} = \{1, \dots, m\} \setminus K$, such that

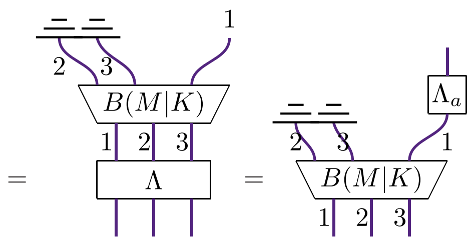


(29)

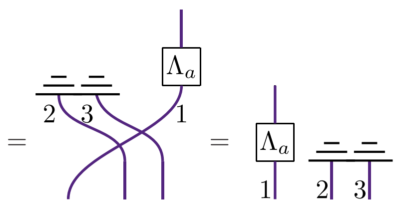
To illustrate this, one of the conditions that this definition would impose on the tripartite case ($M = \{1, 2, 3\}$) would be for $K = \{2, 3\}$, which would give



(30)



(31)



(32)

that is, we can see explicitly how our condition gives us no signaling from $2 \otimes 3$ to 1. It is straightforward to similarly verify that the other conditions in the tripartite case are recovered by varying over the subsets $K \subseteq M$.

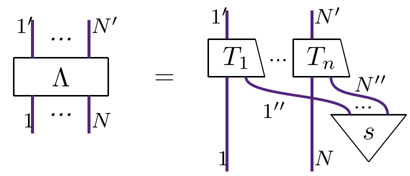
Notice that this definition of a nonsignaling channel treats each pair of input-output systems (i, i') as a different wing of the experiment. Therefore, when specifying the experimental scenario and the channel Λ the systems should be represented via "one wire per wing." As an example, consider the case where one wing of the experiment consists of two qubits forming a four-dimensional quantum system as an input: then this must be represented by one four-dimensional system—rather than by two wires representing two qubits—when Definition III.2 is applied, since signaling is allowed between the wing's internal two qubits.

B. Common cause channels

To formally state the question tackled in this paper, we first need to specify the notion of a common-cause channel that we use in this paper. Broadly speaking, the common-cause channels are a subset of the nonsignaling channels. Namely, we say that a channel is common cause if, in the GPT of interest, it can be constructed by the parties via the application of some local operations to a shared multipartite state. A good example of such a channel is the one obtained in a Bell experiment, where, for example, Alice and Bob each make measurements on their shares of a Bell state. One can view the result of the Bell experiment as being a bipartite classical channel which is realized by local operations on a shared quantum state, i.e., a quantum common cause.

Based on this example, we can define the notion of a common-cause decomposition within a given GPT G.

Definition III.3 (Common-cause decomposition). Let Λ be a channel in a given GPT G. Λ admits of a *common-cause decomposition* if there are N systems $\{1'', \ldots, N''\}$ from G, a state s in the state space of the multipartite system $1'', \ldots, N''$ and a collection $\{T_i\}_{i=1\ldots N}$ of transformations in G, such that



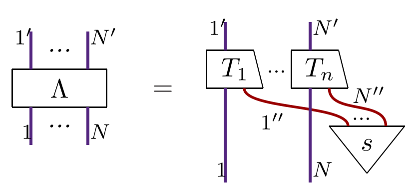
(33)

One can compare this formal diagrammatic definition to the conceptual definition to see that indeed the idea of construction by local operations (the transformations T_i) on a shared common cause (the state s) is indeed captured by this diagram.

Now, the idea of common-cause decomposition within a GPT might not be enough if one is considering the possible existence of some hypothetical cause that might not be modeled by the GPT under consideration. In particular, this is precisely the kind of situation that is considered in the resource theories of Refs. [4,8]. In such cases, the more appropriate question is not whether Λ can be realized with a common cause in G, but whether or not there exists a theory G′ in which it can be realized with a common cause. Going back to our example of the Bell experiment, if we violate a Bell inequality, then we know that the resulting channel cannot be realized via common case within Stoch, but it can be realized via a quantum common cause, that is, within the quantum GPT, Quant.

For that purpose, we define the notion of *GPT-common-cause realizable*, by asking whether the common-cause decomposition of Λ exists in *any* GPT.

Definition III.4 (GPT-Common-cause realizable channel). Let Λ be a channel in a given GPT G. Λ is *GPT-common-cause realizable* if there exists a GPT G′ which contains G as a full subtheory, N systems $\{1'', \ldots, N''\}$ from G′, a state s in the state space of the multipartite system $1'', \ldots, N''$ in G′, and a collection $\{T_i\}_{i=1\ldots N}$ of transformations in G′, such that



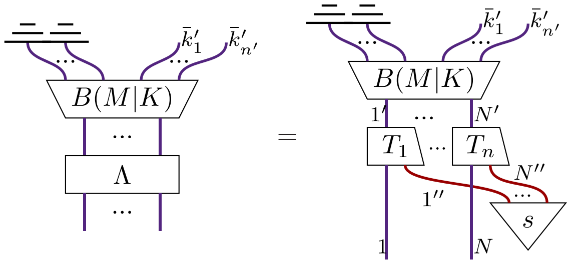
(34)

where we changed the colors of the i'' wires to stress the fact that they can be present only in the hypothetical GPT G′, while the wires i and i' are required to live in the original subtheory G.

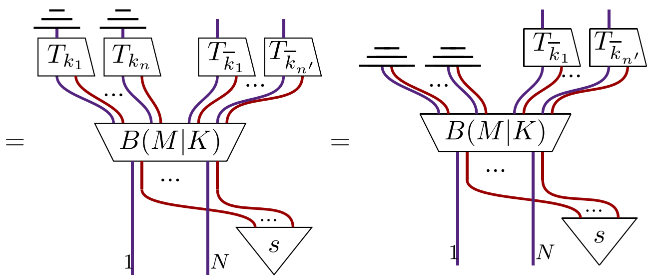
Common-cause realizable channels are well known to be nonsignaling; here we present this result using the diagrammatic notation that we have set up so far.

Proposition III.5. Any GPT-common-cause realizable channel is nonsignaling.

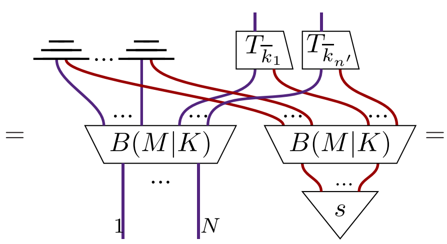
Proof. Consider a fixed but arbitrary channel Λ in a GPT G. Let G′ be the GPT that provides the common-cause realization of Λ. First, notice that, because in Eq. (29) the T_i channels are discard-preserving, if we take Λ to be decomposed as in Eq. (34), we get



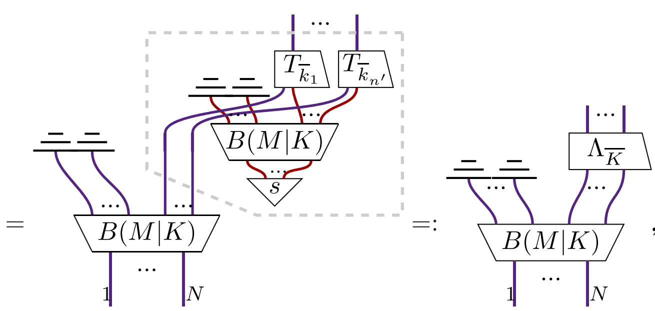
(35)



(36)



(37)



(38)

where $\Lambda_{\overline{K}}$, the channel defined by combining the elements within the dashed box, must be a valid channel from G because all of its inputs and outputs are from G, and G is assumed to be a full subtheory of G′. ■

The main aim of the paper is hence to explore the converse direction to Proposition III.5, namely, whether nonsignaling channels can in general be common-cause realizable. The first observation to make is the well-known fact that the nonsignaling classical channel known [35] as a Popescu-Rohrlich (PR) box [38] does not have a common-cause realization within classical theory [39], but it does have one such realization within the GPT known as Boxworld [16]. In this sense, hence, we say that the classical GPT is *common-cause incomplete*. Moreover, we further view Boxworld as adding extra common causes to classical theory, and so can be thought of as a *common-cause completion* of classical theory. This discussion motivates the following definition:

Definition III.6 (Common-cause complete GPT). A GPT is said to be *common-cause complete* if a common-cause decomposition can be found for each of its nonsignaling channels within the theory. That is, given a nonsignaling channel Λ in the GPT, we can decompose it as in Definition III.4 taking $G' = G$.

The previous observation shows that there are some GPTs—such as classical and quantum theory—which are not common-cause complete. However, classical theory does have a common-cause completion. The question we therefore ask is whether or not this is generic? That is,

Given some GPT G, can we find a common-cause completion G′ such that all of the nonsignaling channels of G have a GPT common-cause realization in G′ (Def. III.4)?

Formally, we defined the common-cause completion as follows:

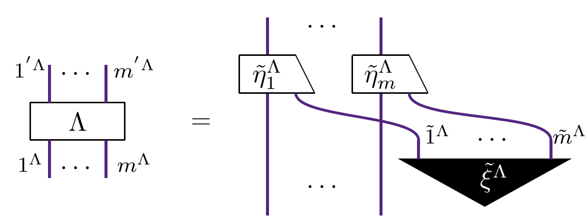
Definition III.7 (Common-cause completion). A GPT G' is a common-cause completion of a GPT G if G is a subtheory of G′, and G′ contains a common-cause decomposition (as per Definition III.4) of all of the nonsignaling channels of G. Note that this definition does not require G′ to be common-cause complete itself.

In the following section we show that any tomographically local GPT does indeed have a common-cause completion.

IV. COMMON-CAUSE COMPLETION

In this section we provide a construction $\mathcal{C}$ which takes an arbitrary locally tomographic causal GPT G into a common-cause completion thereof, $\mathcal{C}[G]$. The starting point of our construction relies on the following lemma, proven in Ref. [40].

Lemma IV.1. (Affine common-cause decomposition of nonsignaling channels [40]) In a locally tomographic GPT G, any m-partite nonsignaling channel, Λ, can be written as



(39)

where $\tilde{\eta}_i^\Lambda$ are discard-preserving processes in G, and $\tilde{\xi}^\Lambda$ is an affine combination of states from G (e.g., when G is quantum theory, $\tilde{\xi}^\Lambda$ is a unit trace Hermitian operator). Note that we have drawn $\tilde{\xi}^\Lambda$ as a black box to indicate that, while it is a mathematically valid object, it is not necessarily a physical process within the GPT G [41].

Proof. Theorem 5.1 of Ref. [40]. ■

This lemma, at first glance, provides a common-cause realization of any nonsignaling channel. However, these affine combinations of states $\tilde{\xi}^\Lambda$ are not (in general) going to be valid states in the GPT. One route to a solution could therefore be to define a common-cause completion by enlarging the state space so that it now includes these nonphysical states. The problem with this approach, however, is that it does not necessarily yield a well-defined GPT, since this procedure will often lead to negative probabilities for measurement outcomes when we start composing these states in ways other than the diagram described in Eq. (39).

In order to prevent negative numbers from arising, then, one can by fiat forbid certain "undesired" compositions. That is, one needs to equip the produced theory with *restrictions* on how the processes may be composed—type-matching conditions would no longer be a sufficient compositional criterion. Such a theory is, in the language of Ref. [42], called a "non-free" GPT as one is not free to compose processes solely based on their system types. While mathematically consistent, we find it difficult to justify such restrictions on physical grounds, and hence we will not pursue its study further in this paper. In what follows, we instead provide a construction of a valid common-cause completion map, which, given a causal tomographically local GPT, will always build a valid GPT, where composition precisely follows the GPT rules as per Sec. II.

A. Constructing the $\mathcal{C}$ map

Here we define a common-cause completion map, $\mathcal{C}$ which takes an arbitrary tomographically local GPT, G, as an input and then constructs a common-cause completion of it, $\mathcal{C}$[G], which is its output. The basic idea of this construction is to include all the nonphysical states $\tilde{\xi}^{\Lambda}$ and $\tilde{\eta}_i^{\Lambda}$ from Lemma IV.1, but now with the caveat that the output systems of each $\tilde{\xi}^{\Lambda}$ (and consequently the inputs to the $\tilde{\eta}_i^{\Lambda}$) are taken to be new system types which are added to the theory. It will then be the type-matching constraints (which are part of the basic definition of a GPT) which will prevent negative probabilities from arising when freely composing processes. It is not immediately clear, however, whether having done so we satisfy all of the other conditions of a GPT, and indeed this turns out not to be the case. Therefore, some extra steps are needed in the construction, in particular, to ensure that the theory is convex and tomographic.

In more detail, the steps followed in the construction, along with what they aim to achieve and how we denote them, are the following:

(1) Take the nonsignaling channels in G and decompose them as per Lemma IV.1. Take each output system of each $\tilde{\xi}^{\Lambda}$ and promote it to a new primitive system type. Collect all these new system types and, together with the system types from G, define a new set of system types including them all. Moreover, include as processes within the theory all of the processes from G together with all processes which are required such that these new systems can realize the common-cause channels as per Lemma IV.1.

Aim: *To ensure that the common-cause decompositions for nonsignaling channels of G exist in* $\mathcal{C}$[G].

Notation: $\mathrm{G} \mapsto \mathrm{G} \sqcup \eta$.

(2) Take the closure of those systems and processes under composition and of the processes under convex combinations.

Aim: *To ensure the compositionality and convexity rules are obeyed.*

Notation: $\mathrm{G} \sqcup \eta \mapsto \mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]$

(3) Quotient the theory $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]$ via operational equivalence.

Aim: *To ensure the theory satisfies tomography.*

Notation: $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}] \mapsto \mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$.

It is this theory that we will define as our common-cause completion, i.e., $\mathcal{C}[\mathrm{G}] := \mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$.

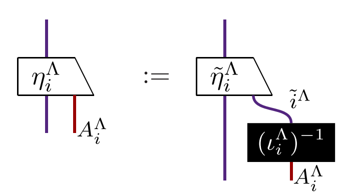
As we progress through the steps, we will show that they do indeed achieve the stated aim. In the end, we will therefore see that the outcome $\mathcal{C}$[G] of this construction is a valid causal GPT (in particular, that there are no extra restrictions on composing systems and processes) and that it is a common-cause completion of G.

In this section we will be dealing with many system types from different GPTs (due to the nature of the problem of extending a theory), and therefore we shall use colors to differentiate the wires corresponding to different theories' system types. The convention we follow is given by the following table:

System type	Wire type
System from the classical subtheory, Stoch	
Generic system from the target GPT, G	
Extra system to be added to G	
Generic system in the new GPT	

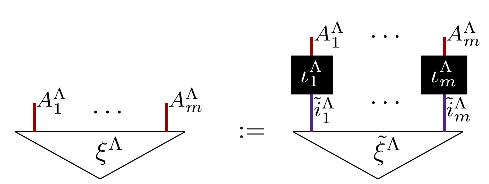
Step 1: Add generating system types and processes

Starting from G, for each Λ in G decomposed as in Eq. (39), let us define a vector space A_i^{Λ} which is isomorphic to $\tilde{i}^{\Lambda}$ with isomorphism $\iota_i^{\Lambda} : \tilde{i}^{\Lambda} \to A_i^{\Lambda}$. Then we define the following linear maps:



(40)

and

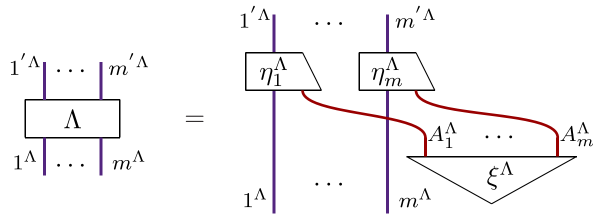


(41)

Note that the isomorphisms ι_i^{Λ} and their inverses are *not* taken to be physically realizable processes within the theory that we are constructing, hence, we denote them, as above, with black boxes. We will, however, take the above composites of them with the $\tilde{\eta}_i^{\Lambda}$ and $\tilde{\xi}^{\Lambda}$, to give η_i^{Λ} and ξ^{Λ}, to be valid processes in the theory we are defining, hence why the left-hand side of Eqs. (40) and (41) are white boxes.

We therefore obtain the following straightforward corollary of Lemma IV.1.

Corollary IV.2. Any m-partite nonsignaling channel, N, can be written as



(42)

Proof. This immediately follows from the definition of the η_i^{Λ} and the ξ^{Λ} [Eqs. (40) and (41)] together with the fact that the ι_i^{Λ} are isomorphisms. ■

We include these extra systems A_i^{Λ} and processes ξ^{Λ}, η_i^{Λ} within the GPT we are building, thereby extending G and enabling the realization of arbitrary nonsignaling channels from G within the common-cause scenario.

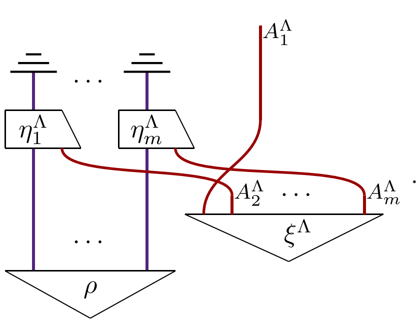
Step 2: Take closure under compositions and convex combinations

For the second step, let us denote by |G| the collection of systems of G, and (with slight abuse of notation) by G the collection of its processes. In order to define the closure properties that we want, we will note that we can view all of the processes that we have defined as living within the process theory of real linear maps, $\mathbb{R}$Linear. To see this, recall that G is, by assumption, tomographically local, and hence is a subtheory of $\mathbb{R}$Linear, and that the new systems and processes that we have added are all, by definition, real linear maps.

We therefore define another subtheory of $\mathbb{R}$Linear which is, by construction, closed under composition as follows.

Definition IV.3. We denote by $\overline{\mathrm{G} \sqcup \eta}$ the subtheory of $\mathbb{R}$*Linear* whose objects (system types) are the closure of $|\mathrm{G}| \sqcup \{A_i^{\Lambda}\}_{\Lambda,i}$ under $\otimes$, and whose morphisms (processes) are the closure of $\mathrm{G} \sqcup \{\eta_i^{\Lambda}, \xi^{\Lambda}\}_{\Lambda,i}$ under $\circ$ and $\otimes$ as the operations in $\mathbb{R}$*Linear*.

Note that, even though we did not explicitly mention the states of the A_i^{Λ} systems, these are implicitly defined by the above closure to obtain $\overline{\mathrm{G} \sqcup \boldsymbol{\eta}}$. For example, by varying over ρ in the following diagram, we can obtain many states of A_1^{Λ}:

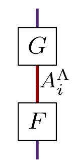


(43)

In the same way, effects and other general processes on the new system types A_i^{Λ} can also be defined. The fact that we have only an implicit definition of the state and effect space is in stark contrast to traditional ways of constructing GPTs, in which the convex geometry of the state and effect spaces is typically the first thing to be defined and then the compositional structure is built on top of this. Here we invert this, first starting with the compositional structure and then defining the geometry of the states and effects which this provides.

Next we will check whether $\overline{G \sqcup \eta}$ leads to sensible probabilistic predictions, namely, whether it contains Stoch as a full subtheory. To answer this we note that Stoch is a full subtheory of G and show that G is a full subtheory of $\overline{\mathrm{G} \sqcup \eta}$, hence, by transitivity, that Stoch is a full subtheory of $\overline{\mathrm{G} \sqcup \eta}$.

Specifically, what we need to show is that any process with all inputs and outputs in |G|, such as



(44)

yields a valid process from G. Note that this is not guaranteed *a priori*, due to the fact that the new systems A_i^{Λ} appear in the interior of the diagram. However, in our case it turns out that this is true as is proven in the following lemma.

Lemma IV.4. Any process in $\overline{\mathrm{G} \sqcup \eta}$ with only input and output system types in |G| is a valid process in G.

Proof. The proof can be found in Appendix B 1. ■

Next we show that $\overline{\mathrm{G} \sqcup \eta}$ is compatible with relativistic causal structure, in the sense that there is a unique effect for each system [29,30].

Lemma IV.5. There is a unique effect for each system in $\overline{\mathrm{G} \sqcup \eta}$.

Proof. The proof can be found in Appendix B 2. ■

A GPT must also be closed under convex combinations so as to model probabilistic mixtures of processes, and so far we have not proven that this is the case for $\overline{\mathrm{G} \sqcup \eta}$. Indeed, it is conceivable that this property has been lost when adding in the new systems and processes and arbitrary diagrams thereof. Hence, we take the convex closure of $\overline{\mathrm{G} \sqcup \eta}$, via the convex combinations of linear maps provided by the supertheory $\mathbb{R}$Linear.

Definition IV.6. We denote by Conv[$\overline{\mathrm{G} \sqcup \eta}$] the convex closure of $\overline{\mathrm{G} \sqcup \eta}$ under convex combinations of processes taken as linear combinations of linear maps from $\mathbb{R}$*Linear*.

Notice that the properties of "has Stoch as a full subtheory" and "is causal" that we proved for $\overline{\mathrm{G} \sqcup \eta}$ are properties which must hold in any GPT, hence we next show they also hold for Conv[$\overline{\mathrm{G} \sqcup \eta}$]:

Lemma IV.7. (1) Any process in Conv[$\overline{\mathrm{G} \sqcup \eta}$] with only input and output system types in |G| is a valid process in G. (2) There is a unique discarding effect for each system in Conv[$\overline{\mathrm{G} \sqcup \eta}$].

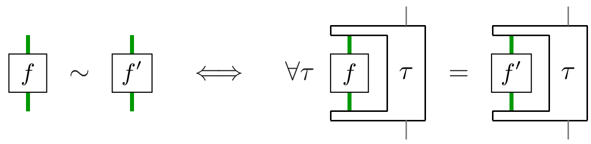
Proof. The proof can be found in Appendix B 3. ■

Step 3: Quotient the theory

There is one final property which must be satisfied in order to have a GPT on our hands, that is, tomography. That means that we need to be able to establish the equality between two processes when the probabilities that they can produce are the same. At this point, however, we do not know that Conv[$\overline{\mathrm{G} \sqcup \eta}$] satisfies this property. Hence, we need a way to "merge" any two differently labeled but operationally equivalent processes (defined shortly) into a single one.

To enforce this, we simply take the quotient $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]$ under operational equivalence. That amounts to defining processes to be equivalence classes and also the operations of sequential, parallel, and convex compositions thereof. For this, let us first formally specify what we mean by "operational equivalence."

Definition IV.8. Processes f and f' (with the same input systems and the same output systems) are operationally equivalent if they give the same statistical predictions when composed with any circuit fragment τ such that the resulting process has only classical inputs and outputs:

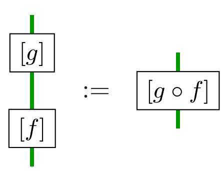


(45)

Note that we are using green wires to denote arbitrary systems which may be G-type, the new systems A_i^{Λ}, or even systems of the quotiented theory, because operational equivalence is a concept defined independently of the theory. In any case, we will apply this here only to $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]$ in order to construct the quotiented theory. We denote the equivalence classes defined by this by square brackets, hence we can write that $f \sim f' \iff [f] = [f']$, and moreover think of some $f' \in [f]$ as providing a representative for the equivalence class of operations that f' belongs to.

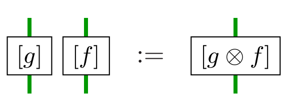
In order to build a theory in which processes are labeled by equivalence classes of processes, we must first define a notion of composition for the equivalence classes.

Definition IV.9. The equivalence classes of processes compose sequentially as



(46)

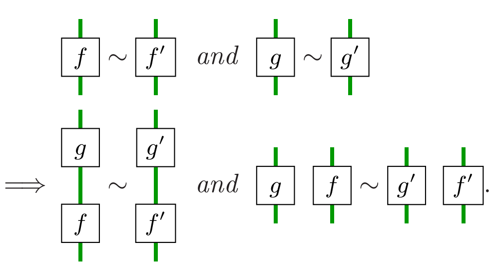
and compose in parallel as



(47)

For these to be valid operations between equivalence classes, they must not depend on the choices of representatives.

Lemma IV.10. Composition as defined in Definition IV.9 is independent of the choices of representatives, i.e.,

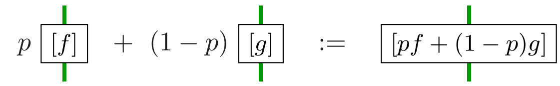


(48)

Proof. The proof can be found in Appendix B 4.

In a similar way we can define convex combinations of equivalence classes as follows:

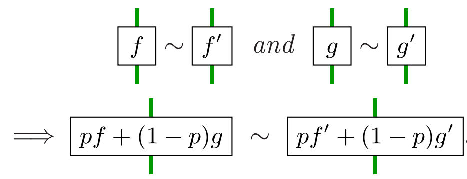
Definition IV.11. Convex mixtures of equivalence classes of processes are given by the following:



(49)

It is easy to see that the relevant properties of convex combinations, for example distributivity over $\circ$ and $\otimes$, are immediately inherited from the analogous property in the prequotiented theory. Again, for consistency, we prove the following:

Lemma IV.12. Convex mixtures as defined in Definition IV.11 are independent of the choice of representative, i.e.,



(50)

Proof. The proof can be found in Appendix B 5. ■

These operations allow us to define the quotiented theory as follows:

Definition IV.13. We denote the theory whose processes are operational equivalence classes of the processes in $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]$, with composition and convex mixtures given by Definitions IV.9 and IV.11, by $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$.

Note that, as G is a GPT, and hence satisfies tomography, for a valid process f in G, we have $[f] = \{f\}$, that is, each equivalence class of processes in G contains a single element. It is then clear that Lemma IV.4 also holds for our quotiented theory. Moreover, it is also clear that Lemma IV.5 continues to hold even in our quotiented theory, as quotienting could identify effects only for a particular system with one another, and as we have only a unique effect for a given system in the first place we have a unique effect after quotienting.

The theory $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$ therefore satisfies all of the desired properties to be considered a causal GPT.

While the GPT that we constructed is $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$, it is clear that it is much easier to perform calculations within $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]$ as it is simply a subtheory of $\mathbb{R}$Linear. Luckily one can always perform calculations in $\mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$ by picking suitable representative elements for the equivalence classes, doing a computation within $\mathbb{R}$Linear, and then requotienting to determine the resultant equivalence class.

Definition IV.14 (Common-cause completion map). The map $\mathcal{C}$ given by $\mathcal{C}[\mathrm{G}] \equiv \mathrm{Conv}[\overline{\mathrm{G} \sqcup \eta}]/\sim$ is a common-cause completion map on the set of causal locally tomographic GPTs.

This is because $\mathcal{C}[\mathrm{G}]$ is a valid GPT which contains G as a full subtheory and where every $\Lambda \in \mathrm{G}$ has a common-cause realization in $\mathcal{C}[\mathrm{G}]$.

V. RESULTS AND DISCUSSION

The construction we have presented for a common-cause completion map is useful as it allows us to understand possible causal explanations of physical phenomena. To elaborate on this, let us first introduce our main theorem and a useful corollary.

Theorem V.1. Given a locally tomographic causal GPT G, its set of multipartite nonsignaling channels (Defition III.2) is the same as its set of multipartite common-cause realizable (Definition III.4) channels. Notice these common causes might not be state preparations allowed in G

Proof. Consider the GPT $\mathcal{C}$[G]. By Proposition III.5, the common-cause realizable channels in G are nonsignaling. In the other direction, by construction, $\mathcal{C}$[G] can provide a common-cause realization of any nonsignaling channel of G. ■

Noting that Quant is a locally tomographic causal GPT we immediately obtain the following corollary.

Corollary V.2. There exists a causal GPT that provides a common-cause realization of every nonsignaling quantum channel. Such a GPT is given by $\mathcal{C}$[Quant].

This corollary is important for two reasons. First, it answers in the negative "Open Question 1" posed in Ref. [8]: *Do there exist bipartite nonsignaling quantum channels which cannot be realized by GPT common causes?*

Second, recall the phenomenon of Einstein-Podolski-Rosen (EPR) inference [43] (a.k.a. steering) where a party (say, Alice) learns about the state preparation of a physical system (held by a distant party, herein called Bob) by performing measurements on her share of the bipartite physical system [44,45]. Here the object of study is the collection of *subnormalized conditional states* that Bob's subsystem may be prepared in, usually called an *assemblage* [46]. Similarly to the case of nonsignaling correlations in Bell experiments, one may mathematically define general assemblages as those which comply with the nonsignaling principle. Given the particular causal structure that underpins these EPR experiments, then, a crucial foundational question is whether these general assemblages could be realized within some (beyond quantum) GPT as a common-cause process. This question can be readily answered in the affirmative by Corollary V.2, given that assemblages in EPR scenarios can be formalized in terms of nonsignaling quantum-classical channels [47]. This sets the foundation stone to be able to study the nonclassicality of EPR assemblages based on the properties of the common-cause process within the GPT that may realize them. In particular, this observation answers in the affirmative the question posed in Ref. [15]: there exists a causal GPT Q′ that provides a common-cause realization of every general assemblage.

More generally, our result provides the fundamental justification of the possibility to assess and quantify the nonclassicality of arbitrary nonsignaling processes by means of the nonclassicality of the common-cause required to realize them. This has previously been argued at length for the case of correlations in Bell scenarios [4], where the existence of common-cause realizations of nonsignaling boxes had already been provided by the GPT known as Boxworld [16]. In this light, hence, our work enables the possibility of extending this causal reasoning to scenarios beyond Bell experiments, which involve other local systems types rather than strictly classical ones.

Looking forward, there are many open research directions pertaining to the common-cause completion construction that we defined:

(1) Is $\mathcal{C}$[G] common-cause complete? Intuitively it seems that this should be the case, but conceivably there may be nonsignaling channels between the new systems which are not realizable in common cause scenarios within $\mathcal{C}$[G]. Note that $\mathcal{C}[\mathcal{C}$[G]] may not be well defined because we do not yet know whether or not:

(a) $\mathcal{C}$[G] is tomographically local or

(b) Whether or not there is a way to extend the common-cause completion to tomographically nonlocal GPTs, or to more general kinds of process theories.

(2) One could also ask questions about properties, such as whether $\mathcal{C}$[G] is the largest or smallest common cause completion of G in some sense; the natural way to answer this question would be by trying to characterize $\mathcal{C}$ via a universal property.

(3) It would also be interesting to give a more concrete definition of $\mathcal{C}$[G] for particular GPTs of interest such as quantum and classical theory. The nature of the construction makes it nontrival to obtain such representations, and, moreover, it does not coincide with the known common cause completion of classical theory known as Boxworld.

While being of technical nature, we expect the answers to these questions to also help us deepen our understanding on the possible nonsignaling processes that can be motivated, understood, and studied from the causal perspective.

ACKNOWLEDGMENTS

P.J.C. and A.B.S. acknowledge support by the Foundation for Polish Science (IRAP project, ICTQT, Contract No. 2018/MAB/5, co-financed by EU within Smart Growth Operational Programme). J.H.S. was supported by the National Science Centre, Poland (Opus project, Categorical Foundations of the Non-Classicality of Nature, Project No. 2021/41/B/ST2/03149). All of the diagrams within this manuscript were prepared using TikZit.

APPENDIX A: (STRICT) SYMMETRIC MONOIDAL CATEGORIES

Since we define a generalized probabilistic theory (GPT) in terms of a strict symmetric monoidal category (SMC), we devote this Appendix to define the latter. We follow that with a brief commentary on interpreting that structure in terms of processes, which is key for understanding how to see that GPTs are SMCs, and end with the most important example of SMC for this paper.

A (strict) symmetric monoidal category consists of (1) a collection of objects $A, B, \ldots$, (2) for each pair of objects A, B, a collection of morphisms $f : A \to B$, and (3) two operations, $\circ$ and $\otimes$, under which the category is closed. The first operation, $\circ$, maps certain pairs of morphisms to morphisms. In particular, it combines $f : A \to B$ and $g : B \to C$ into $g \circ f : A \to C$, and can be performed only when the domain of g matches the codomain of f (in this example, the matching is given by the object B). Furthermore, $\circ$ is associative, so it is similar to function composition. (4) An identity morphism $1_A : A \to A$ that is a unit for $\circ$ is moreover associated with each object A. The second operation, $\otimes$, combines arbitrary pairs of objects, taking A and B to $A \otimes B$ as well as arbitrary pairs of morphisms, taking $f : A \to B$ and $g : C \to D$ into $f \otimes g : A \otimes C \to B \otimes D$. Furthermore, $\otimes$, is associative and has a unit object which we denote I, so it is a monoid operation on the collection of objects, being therefore responsible for the monoidal structure of the category. Finally, the two operations satisfy a consistency condition, namely, that $(g \circ f) \otimes (g' \circ f') = (g \otimes g') \circ (f \otimes f')$.

An interesting property of the symmetric monoidal categories is that they feature a diagrammatic calculus, which provides an intuitive and expressive way to write and perform mathematical calculations. For a description of that, we refer the reader to the Sec. II A.

The bare structure of the SMC has a nice interpretation in terms of processes [17]. We take the objects to represent system types, and call the monoidal unit, denoted I, the trivial system. The morphisms $f : A \to B$ are interpreted as processes that take a system of type A into a system of type B. The processes that start (but do not end) in the unit object (the trivial system), i.e., those like $s : I \to A$, are called states, the ones that end (but do not start) in I, like $e : A \to I$, are called effects, and the ones who neither start nor end in I, such as $f : A \to B$, are called transformations. This is intuitive because $s : I \to A$ can be viewed as some preparation procedure of a system of type A, and $e : A \to I$ as a destructive operation. Next, processes that start and end in I, such as $p : I \to I$, are called *numbers*, or scalars. Now, processes can happen sequentially or in parallel, and this is captured by the SMC: we interpret $g \circ f$ as the sequential composition of the processes f and g, where f is followed by g (which acts on the output of f), and $f \otimes g$ as the composite process given by f and g occurring in parallel. This interpretation of $\circ$ and $\otimes$ motivates the consistency condition that they had to satisfy, since that is the natural relationship between processes happening in parallel and in sequence.

We now illustrate this abstract definition of an SMC by means of the key example for this paper.

Example 1 ($\mathbb{R}$Linear). The SMC $\mathbb{R}$Linear takes objects (system types) to be real vector spaces, and, morphisms (processes) to be linear maps between the vector spaces. The $\circ$ operation is the composition of linear maps, and $\otimes$ is the tensor product. The identity morphisms are given by the identity linear maps, and the monoidal unit is given by the one-dimensional vector space $\mathbb{R}$.

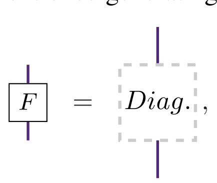
APPENDIX B: PROOFS

1. Proof of Lemma IV.4

Lemma IV.4. Any process in $\overline{\mathrm{G} \sqcup \eta}$ with only input and output system types in $|\mathrm{G}|$ is a valid process in G.

Proof. First note that, by definition, any process in $\overline{\mathrm{G} \sqcup \eta}$ can always be written as a diagram involving only our generating processes, that is, processes in G, and the processes in $\{\eta_i^\Lambda, \xi^\Lambda\}_{\Lambda,i}$.

Now consider an arbitrary process F in $\overline{\mathrm{G} \sqcup \eta}$ with input and output system types in $|\mathrm{G}|$. This process can be written in terms of the above-mentioned generating processes:

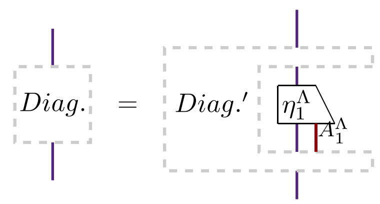


(B1)

where we do not specify the internal structure of the dashed box as the actual compositional structure of F has no generic specification, but we assume it is a diagram consisting of generating processes.

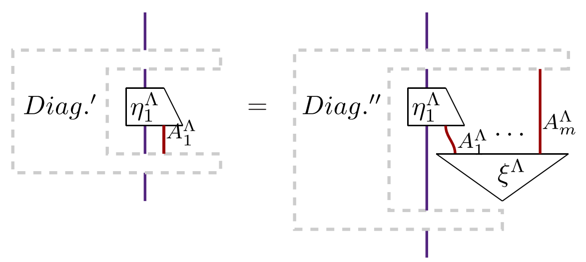
We will now show that this box associated to the process F can always be rewritten into a diagram which involves only processes in G. This follows from the fact the we can rewrite any diagram using only generating processes.

Suppose, for example, that the diagram involves the process η_1^Λ, i.e.,



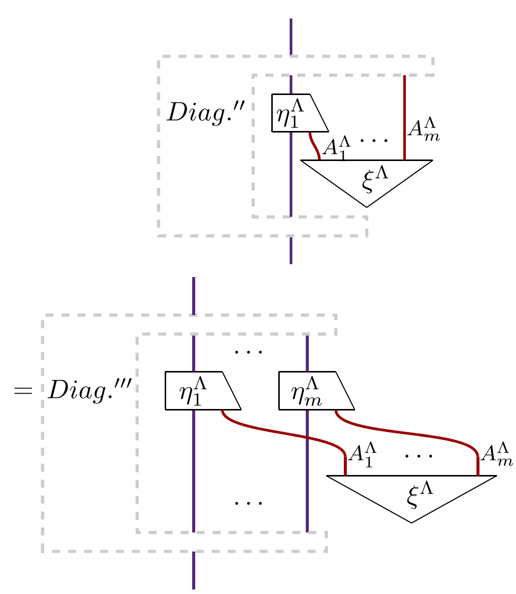
(B2)

Since A_1^Λ is not an input to the process F (as we are assuming that the inputs and outputs system types are in $|\mathrm{G}|$), there must be a process in the diagram $Diag'$ for which this system, A_1^Λ, is an output. There is only one generating process which has A_1^Λ as an output, namely, ξ^Λ. Hence, we can write diagram $Diag'$ as



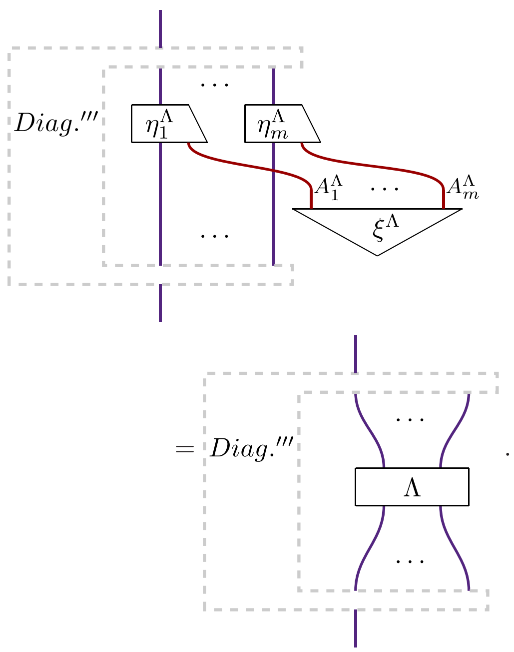
(B3)

We also know that none of the A_i^Λ are outputs of the process, hence, they must be the input of some process within the diagram $Diag''$. For each of these there is a single generating process which has A_i^Λ as an input, namely, η_i^Λ. This means we can rewrite the diagram $Diag''$ as



(B4)

The explicitly drawn part of the diagram, however, is now nothing but the nonsignaling channel Λ, which is a process that lives in G:



(B5)

Hence, we have shown that we can redraw the diagram associated to F so as not to use the generating process η_1^Λ. This argument clearly also applies to any other η_i^Λ that may appear in the specification of F, and a very minor modification of it applies to any ξ^Λ.

This means that any process in $\overline{\mathrm{G}\sqcup\eta}$ whose input and output system has types in |G| can always be written in a way that involves only generating processes from G and does not involve any generating processes from $\{\eta_i^\Lambda, \xi^\Lambda\}$. As G is closed under composition, we have therefore shown that any process with input and output system types in |G| is necessarily a valid process in G. ■

Notice that, in particular, Lemma IV.4 implies that the theory $\overline{\mathrm{G}\sqcup\eta}$ that we have defined will make sensible probabilistic predictions, since the classical systems are valid systems in G and any processes with only classical inputs and outputs is necessarily a stochastic map.

2. Proof of Lemma IV.5

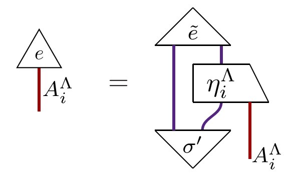
Lemma IV.5. There is a unique discarding effect for each system in $\overline{\mathrm{G}\sqcup\eta}$.

Proof. Here we show that every generating type has a unique discarding effect, as the generalization to composite types is straightforward.

For each generating system i from the GPT G, Lemma IV.4 implies that the discarding effect for i is itself a valid process G. Since G is causal, this means that the discarding effect for i in $\overline{\mathrm{G}\sqcup\eta}$ is unique.

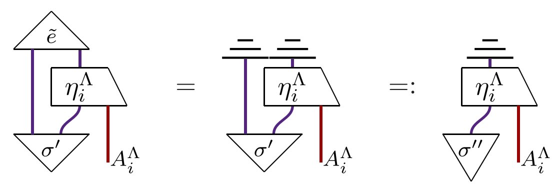
Now we need to show that the claim also holds for system types beyond those present in the GPT G, i.e., the systems $\{A_i^\Lambda\}$.

Since all processes of $\overline{\mathrm{G}\sqcup\eta}$ can decompose in terms of generating processes, we can write a generic effect for A_i^Λ as



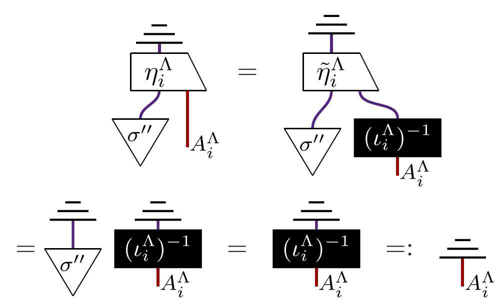
(B6)

where, due to Lemma B 1, we know that $\tilde{e}$ and σ' are necessarily in G. Then, as there is a unique effect for each system in G, we have that



(B7)

Now, using the definition of η_i^Λ we have that



(B8)

Hence the extra systems in the enlarged theory still satisfy the property of having an unique effect to each system type. ■

3. Proof of Lemma IV.7

Lemma IV.7. (1) Any process in Conv[$\overline{\mathrm{G}\sqcup\eta}$] with only input and output system types in |G| is a valid process in G. (2) There is a unique discarding effect for each system in Conv[$\overline{\mathrm{G}\sqcup\eta}$].

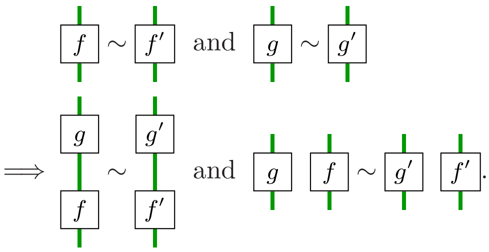
Proof.

(i) From Lemma IV.4 we know that any process in $\overline{\mathrm{G}\sqcup\eta}$ with only input and output system types in |G| is a valid process in G. Since processes in G are closed under convex combinations, this implies that any convex combination of processes in $\overline{\mathrm{G}\sqcup\eta}$ with only input and output system types in |G| is a valid process in G, which proves the claim.

(ii) That there is a unique discarding effect for each system immediately follows from Lemma IV.5 together with the fact that since there is a unique discarding effect for each generating system type it is impossible to obtain other discarding effects by composition and convex combinations. ■

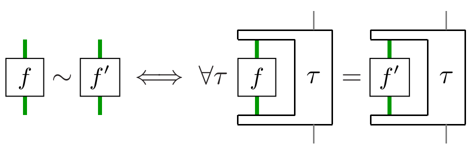
4. Proof of Lemma IV.10

Lemma IV.10. Composition as defined in Definition IV.9 is independent of the choices of representatives, i.e.,



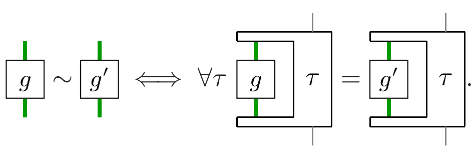
(B9)

Proof. We start by rewriting the assumptions of the theorem using the definition of equivalence:



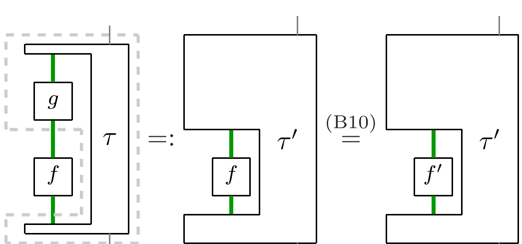
(B10)

and

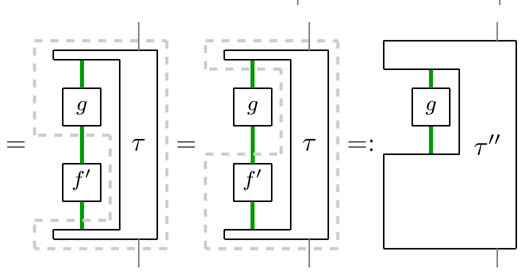


(B11)

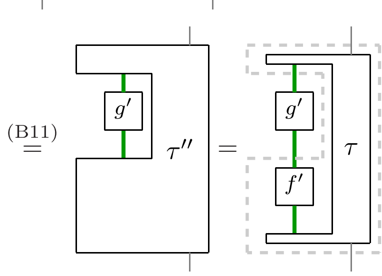
Hence, for an arbitrary τ, the following holds:



(B12)

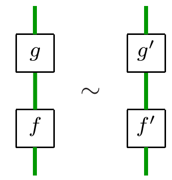


(B13)



(B14)

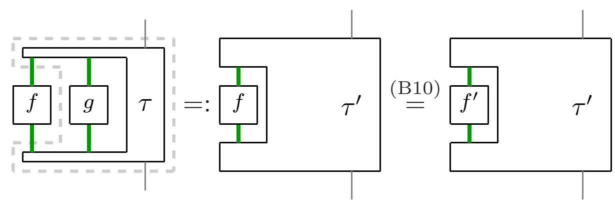
This implies that



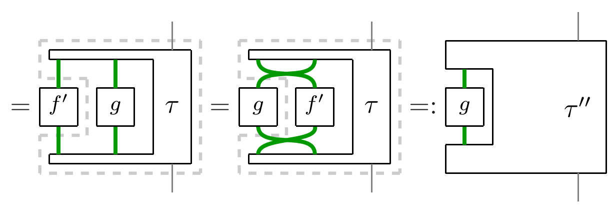
(B15)

hence proving the first part of the lemma.

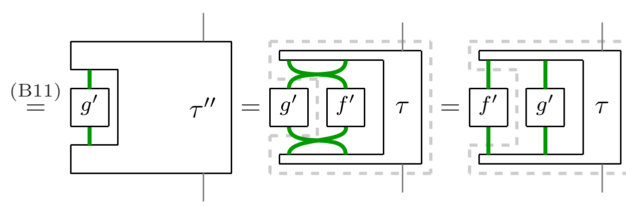
The proof of the second part of the lemma follows in a similar way: for an arbitrary τ,



(B16)

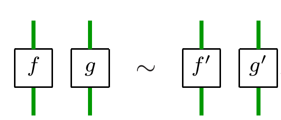


(B17)



(B18)

Hence,

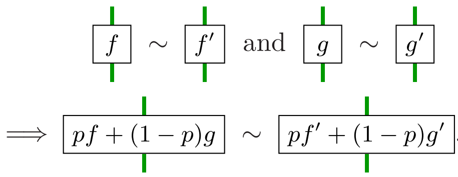


(B19)

which completes the proof. ■

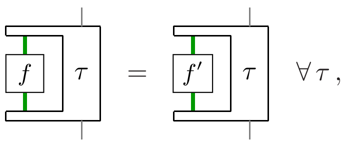
5. Proof of Lemma IV.12

Lemma IV.12. Convex mixtures as defined in Definition IV.11 are independent of the choice of representative, i.e.,



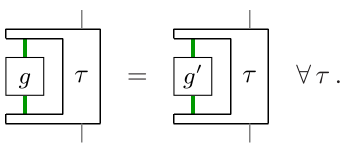
(B20)

Proof. The assumptions of the lemma can be equivalently stated as



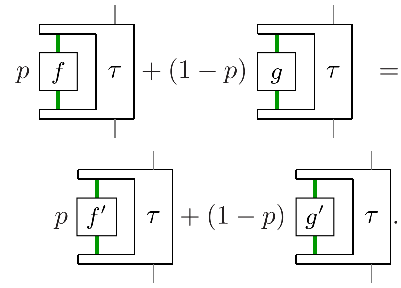
(B21)

and



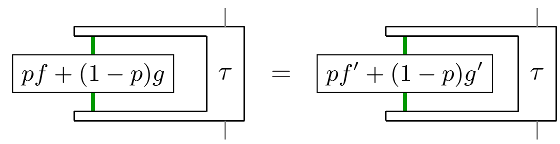
(B22)

In particular, this means that $\forall\, \tau$



(B23)

Linearity of τ implies



(B24)

for all τ, which proves the claim. ■

[1] B. Coecke, T. Fritz, and R. W. Spekkens, Inf. Comput. **250**, 59 (2016).
[2] J. I. de Vicente, J. Phys. A **47**, 424017 (2014).
[3] J. Geller and M. Piani, J. Phys. A **47**, 424030 (2014).
[4] E. Wolfe, D. Schmid, A. B. Sainz, R. Kunjwal, and R. W. Spekkens, Quantum **4**, 280 (2020).
[5] D. Schmid, T. C. Fraser, R. Kunjwal, A. B. Sainz, E. Wolfe, and R. W. Spekkens, Quantum **7**, 1194 (2023).
[6] B. Zjawin, D. Schmid, M. J. Hoban, and A. B. Sainz, Quantum **7**, 926 (2023).
[7] B. Zjawin, D. Schmid, M. J. Hoban, and A. B. Sainz, Quantum **7**, 1134 (2023).
[8] D. Schmid, H. Du, M. Mudassar, G. Coulter-de Wit, D. Rosset, and M. J. Hoban, Quantum **5**, 419 (2021).
[9] D. Schmid, D. Rosset, and F. Buscemi, Quantum **4**, 262 (2020).
[10] D. Rosset, D. Schmid, and F. Buscemi, Phys. Rev. Lett. **125**, 210402 (2020).
[11] C. J. Wood and R. W. Spekkens, New J. Phys. **17**, 033002 (2015).
[12] D. Schmid, J. H. Selby, and R. W. Spekkens, arXiv:2009.03297.
[13] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[14] A. B. Sainz, N. Brunner, D. Cavalcanti, P. Skrzypczyk, and T. Vértesi, Phys. Rev. Lett. **115**, 190403 (2015).
[15] P. J. Cavalcanti, J. H. Selby, J. Sikora, T. D. Galley, and A. B. Sainz, npj Quantum Inf. **8**, 76 (2022).
[16] J. Barrett, Phys. Rev. A **75**, 032304 (2007).
[17] B. Coecke, arXiv:1009.3786.
[18] B. Coecke and A. Kissinger, *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge University Press, Cambridge, UK, 2017).
[19] S. Gogioso and C. M. Scandolo, EPTCS **266**, 367 (2018).
[20] J. H. Selby, C. M. Scandolo, and B. Coecke, Quantum **5**, 445 (2021).
[21] L. Hardy, arXiv:quant-ph/0101012.
[22] M. Müller, SciPost Physics Lecture Notes 028 (2021).
[23] M. Plávala, Phys. Rep. **1033**, 1 (2023).
[24] A. Joyal and R. Street, Adv. Math. **88**, 55 (1991).
[25] P. Selinger, in *New Structures for Physics*, edited by B. Coecke (Springer, 2010), pp. 289–355.
[26] E. Patterson, D. I. Spivak, and D. Vagner, Electron. Proc. Theor. Comput. Sci. **333**, 49 (2021).
[27] L. Hardy and W. K. Wootters, Found. Phys. **42**, 454 (2012).
[28] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).
[29] B. Coecke, Electron. Proc. Theor. Comput. Sci. **172**, 27 (2014).
[30] A. Kissinger, M. Hoban, and B. Coecke, arXiv:1708.04118.
[31] L. Hardy, arXiv:1104.2066.
[32] G. Chiribella, G. M. D'Ariano, and P. Perinotti, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiaribella and R. W. Spekkens (Springer, 2016), pp. 171–221.
[33] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, Quantum **8**, 1283 (2024).
[34] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information: 10th Anniversary Edition* (Cambridge University Press, Cambridge, 2010).
[35] S. Popescu and D. Rohrlich, Found. Phys. **24**, 379 (1994).
[36] D. Beckman, D. Gottesman, M. A. Nielsen, and J. Preskill, Phys. Rev. A **64**, 052309 (2001).
[37] B. Schumacher and M. D. Westmoreland, Quantum Info. Proc. **4**, 13 (2005).
[38] The PR box can be thought of as a nonsignaling classical channel that takes classical systems to classical systems.
[39] J. S. Bell, Phys. Phys. Fiz. **1**, 195 (1964).

[40] P. J. Cavalcanti, J. H. Selby, J. Sikora, and A. B. Sainz, J. Phys. A: Math. Theor. **55**, 404001 (2022).

[41] That is, when we view the tomographically local theory as a subtheory of $\mathbb{R}$Linear we can then take arbitrary affine (or more generally linear) combinations and have a well-defined process in $\mathbb{R}$Linear, but this then could be outside of the subtheory G.

[42] J. Barrett, N. de Beaudrap, M. J. Hoban, and C. M. Lee, npj Quantum Inf. **5**, 41 (2019).

[43] V. P. Rossi, M. J. Hoban, and A. B. Sainz, J. Phys. A: Math. Theor. **55**, 264002 (2022).

[44] E. Schrödinger, Math. Proc. Cambridge Philos. Soc. **32**, 446 (1936).

[45] H. M. Wiseman, S. J. Jones, and A. C. Doherty, Phys. Rev. Lett. **98**, 140402 (2007).

[46] M. F. Pusey, Phys. Rev. A **88**, 032313 (2013).

[47] M. J. Hoban and A. B. Sainz, New J. Phys. **20**, 053048 (2018).