IOP Publishing New J. Phys. 27 (2025) 023004 https://doi.org/10.1088/1367-2630/ada850

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics
Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

CrossMark

OPEN ACCESS

RECEIVED
9 August 2024

REVISED
27 December 2024

ACCEPTED FOR PUBLICATION
9 January 2025

PUBLISHED
4 February 2025

PAPER

Minimal operational theories: classical theories with quantum features

Davide Rolino[1,2,*], Marco Erba[3], Alessandro Tosini[1,2] and Paolo Perinotti[1,2]

1 Dipartimento di Fisica, QUIT Group, Università degli Studi di Pavia, via Bassi 6, Pavia 27100, Italy
2 Sezione di Pavia, INFN Gruppo IV, via Bassi 6, Pavia 27100, Italy
3 International Centre for Theory of Quantum Technologies (ICTQT), Uniwersytet Gdański, ul. Jana Bażyńskiego 1A, Gdańsk 80-309, Poland
* Author to whom any correspondence should be addressed.

E-mail: davide.rolino01@universitadipavia.it, alessandro.tosini@unipv.it, paolo.perinotti@unipv.it and marco.erba@ug.edu.pl

Keywords: operational probabilistic theories, generalised probabilistic theories, operational theories, quantum information, nonclassicality, foundations of quantum theory

Abstract

We introduce a class of probabilistic theories, termed Minimal Strongly Causal Operational Probabilistic Theories, where system dynamics are constrained to the minimal set of operations consistent with the set of states and permitting conditional tests. Specifically, the allowed instruments are limited to those derived from compositions of preparations, measurements, swap transformations, and conditional operations. We demonstrate that minimal theories with conditioning and a spanning set of non-separable states satisfy two quantum no-go theorems: no-information without disturbance and no-broadcasting. As a key example, we construct Minimal Strongly Causal Bilocal Classical Theory, a classical toy-theory that lacks incompatible measurements, preparation uncertainty relations, and is noncontextual (both Kochen–Specker and generalised), yet exhibits irreversibility of measurement disturbance, no-information without disturbance, and no-broadcasting. Therefore, the latter three properties cannot be understood *per se* as signatures of non-classicality. We further explore distinctions between a theory and its minimal strongly causal counterpart, showing that while the minimal strongly causal version of quantum theory diverges from full quantum theory, the same does not hold for classical theory. Additionally, we establish the pairwise independence of the properties of simpliciality, strong causality, and local discriminability.

Contents

1. Introduction

While classical and quantum mechanics appear to be radically different theories, their structure as theories of information processing share many features that can be understood as the minimal features of any theory that aims at describing physical systems and processes. The processing of systems ranges from preparation procedures to evolution and observations using measuring devices. As the puzzle pieces to model any experiment, the above operational primitives have been taken as the basic elements of the framework of Operational Probabilistic Theories (OPTs) [1, 2], then profitably used to understand the origin of quantum peculiarities [3] among all possible theories of information processing. A similar framework is that of Generalised Probabilistic Theories (GPTs) [4–8], which is however focused on statistical aspects arising in prepare-and-measure scenarios, neglecting in most cases the structures related to transformations.

In order to identify a specific theory, it is necessary to make assumptions on the mathematical entities associated with its processes, either directly or by imposing requirements on tasks that can or cannot be actually implemented manipulating the systems of the theory. For example, both classical and quantum states of a composite system, possibly delocalised over separated labs, can be exhaustively probed via local observations in the two laboratories. This natural feature, can instead be violated by theories that are closely related to the quantum one, as the theory of fermionic systems [9, 10], or to the classical one, as in [11] where a non-trivial composition rule for classical systems is introduced. Other assumptions, as the possibility of describing any probabilistic state as part of a perfectly known state of a larger system, sharply discriminate between quantum theory, where it is possible, and classical theory, where it is forbidden. The above criteria are only two instances among several that can in principle bring together or divide classical and quantum theories, or more generally clarify the mathematical structure behind theories of physical systems.

Most studies on OPTs start from the typical assumption that the admissible transformations of each system coincide with the 'maximal set' consistent with the set of states, i. e. the only requirement is that any state must be transformed into another admissible state. This form of *no-restriction hypothesis* leads, for example, to identify the set of quantum transformations with the set of completely positive, trace-preserving maps, and classical transformations with stochastic matrices. In this paper we explore the opposite scenario where one keeps the minimal set of processes compatible with the structures of the framework, given the systems of the theory and their set of states (and of measurements). The latter scenario is closer to that of a real world laboratory, given that experimenters typically do not have access to all theoretically implementable transformations, but only to certain subsets of them.

The resulting class of theories is proved to preserve relevant quantum no-go theorems independently of the nature of its systems. Surprisingly, also classical systems can support quantum features as the impossibility of gaining information without introducing irreversible disturbance, or the impossibility of broadcasting states. This shows that the observation of certain quantum phenomena in an experimental setting cannot always be taken as a definitive proof that the system under study in the experiment is actually quantum[4].

In [13] the class of Minimal Operational Probabilistic Theories (MOPTs) was introduced, namely OPTs where the only allowed operations are preparations, measurements, the identity map, the swap (systems exchange) map, and any operation that can be obtained by composing these operations sequentially or in parallel. A further minimal operation that is missing in MOPTs is the possibility of *conditioning* which experiment to perform next based on the outcome of a previous experiment. This property, also know as *strong causality* [14, 15], is arguably desirable for any reasonable physical theory. Therefore, we introduce a new class of theories, termed Minimal Strongly Causal Operational Probabilistic Theories (MSOPTs), by incorporating all possible conditional operations, while retaining the minimal resource constraints of MOPTs. The main intent is to understand which features of minimal theories are robust under the introduction of strong causality. Remarkably, we show that almost all marking features of minimal models can survive, under suitable conditions—but they can also disappear in other circumstances.

We first systematically analyse the operational *desideratum* that the spaces of operations in an OPT must be 'complete': if there is a procedure to prepare a transformation with arbitrary precision, then the latter is accepted as a transformation of the theory. This property is given in terms of Cauchy sequences with respect to an operational distance for all elements of the theory—states, measurements, transformations and more generally for their collections that generalise the notion of quantum instrument. We introduce a procedure to complete a theory in such a way that it is strongly causal, granting consistency between operational completeness and the compositional structure of the theory.

We then prove that, whenever a minimal theory with strong causality admits a spanning set of non-separable states, the identity transformation for every system is atomic—i.e. it cannot be obtained performing a measurement and ignoring the outcome. This result leads to a series of consequences: MSOPTs satisfy No-Information Without Disturbance (NIWD) [2, 16–18]—that is the impossibility of learning something non-trivial from a system without perturbing its state irreversibly—as well as irreversibility of measurement disturbance [13] and no-broadcasting [5, 19–26].

Moreover, we show that in the above class of theories there exist instances that are classical, in the sense that they are *simplicial* OPTs. In other words, we exhibit theories where the state spaces are simplexes, whose pure states (the vertices of the simplexes) are jointly perfectly discriminable, and at the same time have a wealth of features traditionally considered as signatures of non-classicality. We explicitly construct a classical MSOPT, termed Minimal Strongly Causal Bilocal Classical Theory (MSBCT). This theory is a minimal strongly causal version of Bilocal Classical Theory (BCT) [11]. The present toy-model shares with BCT the property of being locally equivalent to classical theory, thus it has no incompatible measurements, there is no uncertainty in preparation of states and it is noncontextual. However, based on the above results, the model here introduced must exhibit irreversibility, NIWD, and no-broadcasting.

Finally, based on the properties of MSBCT here introduced, we prove the independence of three main features of classical information theory: local tomography, simpliciality of the set of states and strong

[4] A similar result is also discussed in [12]. However, we remark that the two considered scenarios are different. The authors of [12] are interested in assessing the relationship between the classical explainability of parts of theories—in particular, between fragments and shadows of GPTs. While, instead, we consider fully-fledged theories, just with a restriction in the allowed dynamics.

causality. Therefore, our results provide further insights on the relationships between different physical properties: these insights are, in turn, useful for the axiomatisation programs of Quantum Theory (QT), and for adjudicating between QT and alternative physical theories.

The outline of the presentation is as follows. In section 2 a review of the framework of OPTs [1–3, 11, 14, 27] is provided, focusing on those aspects that will be central in this work. Big emphasis, with several original results, is placed on the topological structure that characterises OPTs, introducing the notion of *generalised instrument spaces* (section 2.2.1). In section 3 the properties of Cauchy sequences of transformations and instruments are studied. We show that sequences of instruments are Cauchy if and only if those of the transformations that compose them are (theorem 1), and that the space of deterministic transformations is Cauchy complete (lemma 4). Grounding on the above properties, the procedure for adding all the conditional instruments to a generic OPT is formalised (section 4). Theorem 2 guarantees that adding the conditional operations and subsequently Cauchy completing the transformations spaces is sufficient for the new theory obtained to be strongly causal. After recalling the definitions of the properties of interest for this article (broadcasting and irreversibility of measurement disturbance) (section 5), in section 6 the classes MOPTs and MSOPTs (section 7) are introduced. We show that in MOPTs the identity is atomic for every systems (theorem 6), while in MSOPTs this property holds for every theory where the entangled states are spanning for the set of states (theorem 7). Then, in section 8, we construct MSBCT, characterizing its main properties. In conclusion, an in-depth discussion of the consequences of the results is carried out (section 10).

2. Operational Probabilistic Theories

In the last two decades, the way we study and understand the quantum world has profoundly changed. With the advent of quantum information [4, 28–31] we started to treat QT as a theory of information processing [6, 31, 32] selected among a Universe of possible alternative theories [1–4, 7, 27, 33, 34]. The selection criteria pertain to the ability to perform specific information processing tasks [2, 3].

Aim of the framework of operational probabilistic theories (OPTs) is exactly to model QT along with all these other alternative theories of information processing and to describe every information theory starting from its compositional (combining operations to build up experiments) structure. The same aim and scope is shared with the deeply related frameworks of GPTs [6, 35–37] and quantum picturialism [38, 39], with common roots dating back to Ludwig's works on the foundations of quantum mechanics [40, 41].

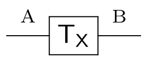
In this section we provide a review of the framework of OPTs [1, 2, 14, 27, 42] emphasizing their *linear* and *topological* structure and proving some properties later used in this work.

2.1. Basic structure of the framework

Every OPT Θ is completely characterised by a set of systems along with the set of operations that it is possible to perform on them.

Systems represent the physical entities which are probed in a laboratory (e.g. an electron, a molecule, a radiation field, etc...) [42]. They are denoted with capital Roman letters $\mathrm{A}, \mathrm{B}, \ldots \in \mathsf{Sys}(\Theta)$. In QT systems are complex Hilbert spaces. The processes occurring between systems are captured by the notion of *tests*, which represent physical processes that can occur within a given theory. A given test $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \equiv \mathsf{T}_{\mathsf{X}} \in \mathsf{Test}(\mathrm{A}\to\mathrm{B})$ models an experiment acting on a given input system A with output system B[5]. Systems can also be depicted in diagrammatic notation as wires, while tests as wired boxes:

$$\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}} \quad \longleftrightarrow$$



As a convention the input–output direction is taken to go from the left to the right, which does not imply a preferred direction for the flow of information[6].

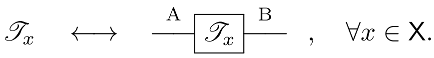
The 'X' appearing in the definition of a test represents the *outcome space* of the test. It is a finite set containing all the possible outcomes of the experiment. To each outcome $x \in \mathsf{X}$ is associated an *event* $\mathscr{T}_x \in \mathsf{Event}(\mathrm{A}\to\mathrm{B})$ representing the realization of a particular occurrence in a physical process[7]. Therefore,

[5] Most of the times, unless it is not clear from the context, the input and output systems of a test will not be specified, thus preferring the notation T_{X} in place of $\mathsf{T}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{B}}$.

[6] In the subset of *causal OPTs* a preferred direction for the flow of information is instead fixed, typically from the left to the right indeed.

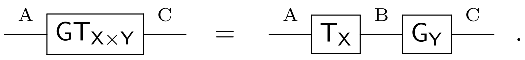
[7] Here we do not include the possibility of having infinite, possibly continuous, outcome spaces. However, the framework has no bottlenecks towards non-finite outcome spaces.

tests are finite collections of events: $\mathsf{T}_{\mathsf{X}} \equiv \{\mathscr{T}_x\}_{x\in\mathsf{X}}$. Diagrammatically:

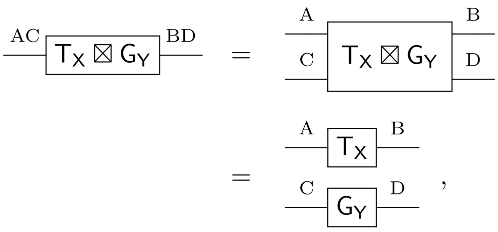


There exists a particular set of events called *deterministic* which are the ones associated to tests whose outcome space has just one element—that is a *singleton set*—, which will be represented as $\star := \{*\}$. Tests associated to deterministic events are called *singleton tests* and operationally model processes that do not provide information. In QT, tests are *quantum instruments*, events are *quantum operations*, and deterministic events are *quantum channels*.

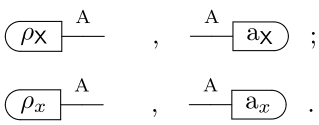
Tests, and consequently events, can be composed in two ways. Sequentially:



and in parallel:

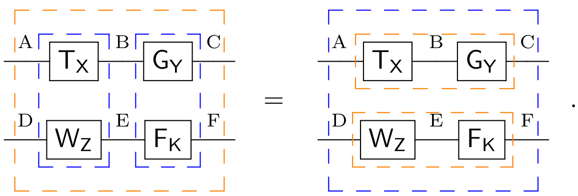


where AB is a *composite system*, obtained by composing in parallel the two systems A and B. The operation of parallel composition of systems is associative and has an identity element: the trivial system. The trivial system I is a particular system representing 'nothing the theory cares to describe' [27]. Furthermore, the set of systems $\mathsf{Sys}(\Theta)$ of any OPT is closed with respect to the latter operation[8]. In OPTs, in general, the operation of parallel composition $\boxtimes$ differs from the standard tensor product, as it happens for example in the composition of fermionic systems in Fermionic quantum theory [9, 10, 43–45] or in classical theory with bilocal tomography [11]. Both operations of sequential and parallel composition are associative and have an identity element. In the former case the identities are given by a family of tests $\{\mathsf{I}_\star^{\mathrm{A}\to\mathrm{A}}\}_{\mathrm{A}\in\mathsf{Sys}(\Theta)}$ with the associated family of deterministic events $\{\mathscr{I}_{\mathrm{A}}\}_{\mathrm{A}\in\mathsf{Sys}(\Theta)}$, while in the latter case the identity is given by $\mathscr{I}_{\mathrm{I}}$. F Diagrammatically the trivial system will not be represented, leaving a blank space. Accordingly tests of the form $\rho_{\mathsf{X}}^{\mathrm{A}\to\mathrm{I}} \in \mathsf{Test}(\mathrm{I}\to\mathrm{A})$ and $\mathsf{a}_{\mathsf{X}}^{\mathrm{A}\to\mathrm{I}} \in \mathsf{Test}(\mathrm{A}\to\mathrm{I})$, and the corresponding events, will be represented as follows:



Tests of this kind are called *preparation-* and *observation-tests* of system A, respectively, while their associated events are called *preparations* and *observations* of system A, respectively. These represent a generalisation of the notions of density matrix and Positive Operator-Valued Measure (POVM) of QT, respectively. When writing the equations not in diagrammatic form, the *round ket* $|\cdot)$ and *round bra* $(\cdot|$ notation will be used to represent preparations and observations, respectively. The last particular case we have to consider is where both the input and output systems are the trivial one. In this case, the tests $p_{\mathsf{X}} \in \mathsf{Test}(\mathrm{I}\to\mathrm{I})$ are called *scalar-tests*, while the corresponding events $p_x \in \mathsf{Event}(\mathrm{I}\to\mathrm{I})$ are called *scalars*.

In conclusion, we observe that the two operations of parallel and sequential composition are required to commute:



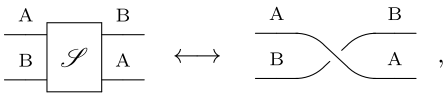
(1)

[8] The fact that $\mathsf{Sys}(\Theta)$ is closed with respect to the operation of parallel composition means that if any two systems A and B are in $\mathsf{Sys}(\Theta)$, then also $\mathrm{AB} \in \mathsf{Sys}(\Theta)$.

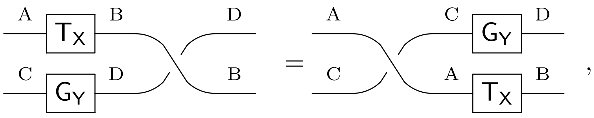
Before concluding our presentation of the compositional structure of OPTs, we introduce the notion of *reversible* event. An event $\mathscr{R} \in \mathsf{Event}(\mathrm{A} \to \mathrm{B})$ is *reversible* if there exists an event $\mathscr{R}^{-1} \in \mathsf{Event}(\mathrm{B} \to \mathrm{A})$ such that $\mathscr{R} \circ \mathscr{R}^{-1} = \mathscr{I}_{\mathrm{A}}$ and $\mathscr{R}^{-1} \circ \mathscr{R} = \mathscr{I}_{\mathrm{B}}$.

The introduction of reversible transformations also allows to define the notion of *operational equivalence* between systems. Two systems A and B are said to be operationally equivalent $\mathrm{A} \cong \mathrm{B}$ if there exists a reversible transformation $\mathscr{R} \in \mathsf{RevTransf}(\mathrm{A} \to \mathrm{B})$, i.e. whose input and output systems are A and B or viceversa.

We are now able to make the final requirement so that the compositional structure of OPTs turns out to be analogous to that of QT. We require the existence of a family of reversible tests called *braiding* which allows a pair of agents to exchange systems between each other. In other words, given any two systems A, $\mathrm{B} \in \mathsf{Sys}(\Theta)$ there exist two singleton reversible tests: $\mathsf{S}_{\star}^{\mathrm{AB} \to \mathrm{BA}} = \{\mathscr{S}_{\mathrm{A,B}}\}$ and its inverse $(S^{-1})_{\star}^{\mathrm{BA} \to \mathrm{AB}} = \left\{\mathscr{S}_{\mathrm{A,B}}^{-1}\right\}$. They will be pictorially represented as follows:

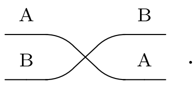


These tests must satisfy the *naturality* property:



(2)

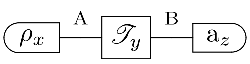
namely tests and events can slide along the wires. OPTs where $\mathscr{S}_{\mathrm{A,B}} = \mathscr{S}_{\mathrm{A,B}}^{-1}$ for any pair of systems of the theory are called *symmetric*. In this case the braiding operation becomes a transposition (also referred to as the *swap* operation) and it is diagrammatically represented as follows:



The described structure is that of a braided strict monoidal category [46–48] and carries only an operational interpretation. It is merely a descriptive tool. To allow OPTs to make predictions about experiments' outcomes, we have to supply them with a *probabilistic structure*. It is then required that to any acyclic circuit of events beginning with a preparation and ending with an observation, i.e. a scalar event, it is associated a conditional probability distribution

$$\mathsf{Event}(\mathrm{I} \to \mathrm{I}) \ni \mathrm{G}\left(\mathscr{T}_x, \mathscr{G}_y, \ldots, \mathscr{F}_z\right)$$
$$:= \mathbb{P}\left(x, y, \ldots, z | \mathrm{G}\left(\mathsf{T}_\mathsf{X}, \mathsf{G}_\mathsf{Y}, \ldots, \mathsf{F}_\mathsf{Z}\right)\right).$$

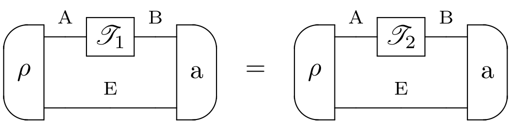
In other words, given that the experiment $\mathrm{G}\left(\mathsf{T}_\mathsf{X}, \mathsf{G}_\mathsf{Y}, \ldots, \mathsf{F}_\mathsf{Z}\right)$ is performed, the formula provides the probability of occurrence of any series of events $\mathrm{G}\left(\mathscr{T}_x, \mathscr{G}_y, \ldots, \mathscr{F}_z\right)$ and then reading the corresponding series of outcomes $(x, y, \ldots, z)$. For example:



$$:= p\left(x, y, z | \rho_\mathsf{X}, \mathsf{T}_\mathsf{Y}, \mathrm{a}_\mathsf{Z}\right),$$

where ρ_x is an event of the test ρ_X and analogously for the others.

Within the framework it is also made the requirement that the spaces of tests and events are quotiented with respect to the following equivalence relation. For all systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and for all events $\mathscr{T}_1$, $\mathscr{T}_2 \in \mathsf{Event}(\mathrm{A} \to \mathrm{B})$, we define $\mathscr{T}_1 \sim \mathscr{T}_2$ if and only if



(3)

for all possible $\mathrm{E} \in \mathsf{Sys}(\Theta)$, $\rho \in \mathsf{Event}(\mathrm{I} \to \mathrm{AE})$, and $\mathrm{a} \in \mathsf{Event}(\mathrm{BE} \to \mathrm{I})$. This follows from the idea that whenever two events (tests) are characterised by the same statistics in any experiment they are indistinguishable. We observe that, while it is true that whenever $\mathscr{T}_1 = \mathscr{T}_2$ also (3) also holds, the converse is not true in general, hence the requested equivalence relation.

The quotient class of events in an OPT are called *transformations*, and their subset having input system A and output system B is denoted by:

$$\mathsf{Transf}(\mathrm{A} \to \mathrm{B}) := \mathsf{Event}(\mathrm{A} \to \mathrm{B})/_{\sim}.$$

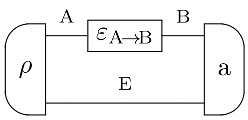
The special case of preparations $\mathsf{St}(\mathrm{A}) := \mathsf{Transf}(\mathrm{I} \to \mathrm{A})$ and observations $\mathsf{Eff}(\mathrm{A}) := \mathsf{Transf}(\mathrm{A} \to \mathrm{I})$ are called the *states* and *effects* of system A, respectively. We define $\mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}$, $\mathsf{St}_1(\mathrm{A})$ and $\mathsf{Eff}_1(\mathrm{A})$ as the set of deterministic transformations, states and effects, respectively. The tests from A to B become *instruments*, and their collection is denoted by:

$$\mathsf{Instr}(\mathrm{A} \to \mathrm{B}) := \mathsf{Test}(\mathrm{A} \to \mathrm{B})/_{\sim}.$$

Finally, the collections of *preparation*- and *observation-instruments* are denoted by $\mathsf{Prep}(\mathrm{A})$ and $\mathsf{Obs}(\mathrm{A})$, respectively.

There is one final assumption for a generic information theory to be an OPT: the theory must be closed with respect to the *coarse-graining* operation. This operation allows one to disregard information related to the outcome of an experiment. Given any test T_X and any disjoint partition $\{\mathsf{Z}^{(y)}\}_{y\in\mathsf{Y}}$ of the outcome space X there exists the *coarse-grained test* T'_Y representing the same operation, where the outcome $y \in \mathsf{Y}$ stands for 'the outcome of the test T_X belongs to $\mathsf{Z}^{(y)}$'. The event $\mathscr{T}_y{}' = \sum_{x\in\mathsf{Z}^{(y)}} \mathscr{T}_x$ is called *coarse-grained*. Obviously, given a test T_X the full coarse-grained transformation $\mathscr{T}_\mathsf{X} = \sum_{x\in\mathsf{X}} \mathscr{T}_x$ is deterministic. The operations of scalar multiplication, sequential and parallel composition distribute over coarse-graining. As to what precisely is meant by the sum symbol used here we refer to section 2.2.

In conclusion of this section, we observe that for every OPT the event with null-probability in any experiment is an actual event of the theory: $0 \in \mathsf{Transf}(\mathrm{I})$. Consequence of this fact is that for every pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, there exists a transformation $\varepsilon_{\mathrm{A}\to\mathrm{B}} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$, called *null transformation*, defined by the following relation:



$$= (\mathrm{a}|_{\mathrm{BE}} \, (\varepsilon_{\mathrm{A}\to\mathrm{B}} \boxtimes \mathscr{I}_{\mathrm{E}}) \, |\rho)_{\mathrm{AE}} = 0,$$

for any system $\mathrm{E} \in \mathsf{Sys}(\Theta)$, $\rho \in \mathsf{St}(\mathrm{AE})$, and $\mathrm{a} \in \mathsf{Eff}(\mathrm{BE})$. In words, the null transformation is the transformation that always occur with null probability in any closed circuit. Clearly, transformations are invariant for coarse-graining with the null one

$$\mathscr{T} = \mathscr{T} + \varepsilon_{\mathrm{A}\to\mathrm{B}},$$

for any couple of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$.

2.1.1. Atomicity, extremality and purity

We introduce a classification of transformations based on how they can be decomposed as combinations of other transformations of the theory.

Definition 1 (Atomic transformation). A transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is *atomic* if, given $\mathscr{T}_1, \mathscr{T}_2 \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$, one has the following implication:

$$\mathscr{T} = \mathscr{T}_1 + \mathscr{T}_2 \implies \mathscr{T}_1, \mathscr{T}_2 \propto \mathscr{T}.$$

The notion of atomic transformations captures the idea of 'indecomposable' events from a conic point of view. These are the transformations that generate the extremal rays of the cones generated by the transformation sets:

$$\mathsf{Transf}_{+(\mathrm{A}\to\mathrm{B})} := \{\lambda\mathscr{T} \mid \lambda \geqslant 0, \mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})\}.$$

The same argument can also be made in the case of convex combinations.

Definition 2 (Extremal transformations). A transformation $\mathscr{T} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is called *extremal* if, given $\mathscr{T}_1, \mathscr{T}_2 \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and $p \in (0,1)$, the condition $\mathscr{T} = p\mathscr{T}_1 + (1-p)\mathscr{T}_2$ implies $\mathscr{T}_1 = \mathscr{T}_2 = \mathscr{T}$.

Extremal transformations embody the notion of extreme points of convex sets. The latter property is of particular interest in the special class of theories where $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is convex for every A and B

Definition 3 (Convex OPTs). An OPT Θ where $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ coincides with its convex hull for every couple of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ is called *convex* [2, 11].

In general the two properties of atomicity and extremality of a transformation are not related. There are transformations that are extremal but not atomic and viceversa, an example being the deterministic effect of any system of Classical Theory (CT) or QT [2], which is clearly an extremal point of the convex set of effects but can be obtained as the coarse-graining of any observation-test.

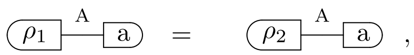
Slightly detaching ourselves from QT's tradition, and following the nomenclature of [11], we have the following definition.

Definition 4 (Pure and mixed transformations). A transformation is *pure* if it is extremal and deterministic. While, a transformation is *mixed* if it is neither atomic, nor extremal.

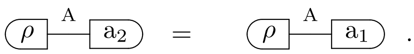
In the following, the set of pure states of a system A will be indicated with $\mathsf{PurSt}(\mathrm{A})$.

2.2. Linear structure

The equivalence relation of transformation (3) reduces to the following in the case of states:



and analogously for effects:



These relations have two important consequences. First, the set of states is separating for that of effects—that is, for every pair of states $\rho_1, \rho_2 \in \mathsf{St}(\mathrm{A})$ such that $\rho_1 \neq \rho_2$, there exists an effect $\mathrm{a} \in \mathsf{Eff}(\mathrm{A})$ such that $(\mathrm{a}|\,\rho_1)_\mathrm{A} \neq (\mathrm{a}|\,\rho_2)_\mathrm{A}$—and viceversa for effects. Second, states can be seen as a set of functionals from $\mathsf{Eff}(\mathrm{A})$ to the real interval $[0,1]$, and viceversa effects are a set of functionals that map $\mathsf{St}(\mathrm{A})$ to $[0,1]$.

These two properties allows us to equip OPTs with a linear structure by extending the functional described above to the whole $\mathbb{R}$ [2]. Considering the collection of all the functionals defined in this way it is possible to construct the spaces of *generalised states* $\mathsf{St}_\mathbb{R}(\mathrm{A})$ and *generalised effects* $\mathsf{Eff}_\mathbb{R}(\mathrm{A})$, which are the real vector spaces for which $\mathsf{St}(\mathrm{A})$ and $\mathsf{Eff}(\mathrm{A})$ are spanning sets, respectively. The separability property between states and effects induces the same property for the generalised spaces. Furthermore, each one is included in the algebraic dual of the other, $\mathsf{St}_\mathbb{R}(\mathrm{A}) \subseteq \mathsf{Eff}_\mathbb{R}(\mathrm{A})^\vee$ and $\mathsf{Eff}_\mathbb{R}(\mathrm{A}) \subseteq \mathsf{St}_\mathbb{R}(\mathrm{A})^\vee$. In the particular case where the dimension $\dim \mathsf{St}_\mathbb{R}(\mathrm{A})$ (or equivalently $\dim \mathsf{Eff}_\mathbb{R}(\mathrm{A})$) is finite one has $\mathsf{St}_\mathbb{R}(\mathrm{A}) = \mathsf{Eff}_\mathbb{R}(\mathrm{A})^\vee$ ($\mathsf{Eff}_\mathbb{R}(\mathrm{A}) = \mathsf{St}_\mathbb{R}(\mathrm{A})^\vee$). The dimension of the generalised state space $\mathsf{D}_\mathrm{A} := \dim \mathsf{St}_\mathbb{R}(\mathrm{A})$ is defined to be the *dimension* (or *size*) of system A. The dimension D of a system represents the number of probabilities that has to be known in order to completely characterise the states of the system when represented as vectors in $\mathbb{R}^\mathsf{D}$. For example, in QT the size of a system A defined on an Hilbert space of dimension d_A is given by $\mathsf{D}_\mathrm{A} = d_\mathrm{A}^2$.

It is also possible to define *generalised transformations* $\mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$. Looking at (3), one has to consider the families of transformations $\{\mathscr{T} \boxtimes \mathscr{I}_\mathrm{E}\}_{\mathrm{E}\in\mathsf{Sys}(\Theta)}$ seen as maps between the collections of states $\mathsf{St}(\mathrm{AE})$ and $\mathsf{St}(\mathrm{BE})$, or through their dual between the collections of effects $\mathsf{Eff}(\mathrm{BE})$ and $\mathsf{Eff}(\mathrm{AE})$. Starting from them, one can define a unique family of linear maps between the generalised spaces and consequently construct the real vector space $\mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$. The operations of parallel and sequential composition for the generalised case are induced by the same operations for transformations [2].

We observe that the summation symbol used to indicate the coarse-graining operation has now a precise meaning. Whenever, we make the coarse-graining of two operations, we are considering their sum, seeing the transformations as elements of the generalised transformation space.

The linear structure makes any OPT 'usable'. In fact, regardless of how abstract are its operational constituents it is always possible to embed everything in a linear vector space where calculations can be made.

As already observed, in OPTs the rule for parallel composition is normally not given by the standard tensor product $\otimes$. Therefore, in general it holds that $\mathsf{St}_\mathbb{R}(\mathrm{A}) \otimes \mathsf{St}_\mathbb{R}(\mathrm{B}) \subseteq \mathsf{St}_\mathbb{R}(\mathrm{AB})$ (and similarly, $\mathsf{Eff}_\mathbb{R}(\mathrm{A}) \otimes \mathsf{Eff}_\mathbb{R}(\mathrm{B}) \subseteq \mathsf{Eff}_\mathbb{R}(\mathrm{AB})$).

2.2.1. Generalised instruments

The analysis of the linear structure of instruments, which is relevant for the present purposes, is less straightforward.

Instruments are a family of finite ordered n-tuples of transformations with constraints for the compatibility of the general structure of the theory. For example, the full coarse-graining of an instrument must be a deterministic transformation, or the sequential composition of two instruments must be another instrument. Indeed, within the framework of OPTs the most elementary concept is that of instrument, and not that of transformation: the set of instruments is defined first and the set of transformations follows. It is always true that

$$\mathsf{Instr}(\mathrm{A}\to\mathrm{B}) \in \mathscr{P}_{\mathrm{ord}}\{\mathsf{Transf}(\mathrm{A}\to\mathrm{B})\}$$

for all $\mathrm{A},\mathrm{B}\in\mathsf{Sys}(\Theta)$, where $\mathscr{P}_{\mathrm{ord}}\{S\}$ represents the set of all ordered subsets of S. Therefore, the most natural way to define a sort of generalised space for instruments is:

$$\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B}) := \bigoplus_{n\in N}\mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}, \tag{4}$$

where N is the cardinality of the outcome space of the generalised instruments. We recall that by hypothesis the cardinality of the outcome space of instruments is finite, and the operation of direct sum is always well defined. $\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ is still a vector space for any $N\in\mathbb{N}$ with the usual scalar multiplication:

$$p(\mathscr{T}_1,\dots,\mathscr{T}_N) = (p\mathscr{T}_1,\dots,p\mathscr{T}_N),$$

and elementwise sum:

$$(\mathscr{T}_1,\dots,\mathscr{T}_N)+(\mathscr{G}_1,\dots,\mathscr{G}_N) = (\mathscr{T}_1+\mathscr{G}_1,\dots,\mathscr{T}_N+\mathscr{G}_N),$$

where the scalar multiplication $p\mathscr{T}_1$, with $p\in\mathsf{Transf}_{\mathbb{R}(\mathrm{I})}$, and the sum operation $\mathscr{T}_1+\mathscr{G}_1$ are the ones of $\mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$.

The generalised spaces of instruments have the interesting property that $\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ can always be seen as a subspace of $\mathsf{Instr}^{(M)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ whenever $N<M$. This comes from the fact that an instrument with N outcomes can always be seen as an instrument with M outcomes of which $M-N$ occur with zero probability, i.e.

$$\{\mathscr{T}_x\}_{x\in\mathsf{X}} \to \{\mathscr{T}_x\}_{x\in\mathsf{X}} \bigcup_{n\in(M-N)}\{\varepsilon_{\mathrm{A}\to\mathrm{B}}\}, \tag{5}$$

where $\{\mathscr{T}_x\}_{x\in\mathsf{X}}\in\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$.

The coarse-graining operation on these generalised instruments is simply defined by extension through linearity from the coarse-graining defined on the allowed instruments of a theory. In practice the coarse-graining operation can be seen as a linear function from $\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ into $\mathsf{Instr}^{(M)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ with $M<N$.

2.3. Topological structure and operational completeness

Here we show how to induce a *topological structure* on the generalised transformation space using the *operational norm* $\|\cdot\|_{op}$ [2]. The distance, induced by the norm, is related to the probability of discrimination between two transformations; the closer they are, the harder it is to discriminate between them. The operational norm allows one to express the requirement that the spaces of transformations, states, and effects are Cauchy complete, corresponding to the operational *desideratum* that if there is a procedure to prepare a transformation with arbitrary precision, then it is natural to assume that the latter is a transformation of the theory.

The operational norm is not the only norm one can choose for an OPT. Another norm of interest is the *sup norm* $\|\cdot\|_{sup}$, used for example to construction quasi-local algebras of transformations in the OPTs framework [14]. In general, these two norms are not equivalent and satisfy different properties. However, in the following discussion we will restrict to the class of OPTs where these two norms are equivalent, in particular, focusing on OPTs where all spaces of instruments, and consequently transformations, are finite dimensional.

Before proceeding we just state some properties of the aforementioned norms, since they will be used throughout our argument.

Lemma 1 (Monotonicity of the operational norm). *Let* $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}(\mathrm{B}\to\mathrm{C})}$, *then*

$$\|\mathscr{T}\|_{op} \geqslant \|\mathscr{E}\mathscr{T}\mathscr{C}\|_{op}, \tag{6}$$

where $\mathscr{E} \in \mathsf{Transf}_{1(\mathrm{C}\to\mathrm{D})}$, *and* $\mathscr{C} \in \mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}$. *The equality holds if both* $\mathscr{E}$ *and* $\mathscr{C}$ *are reversible [2].*

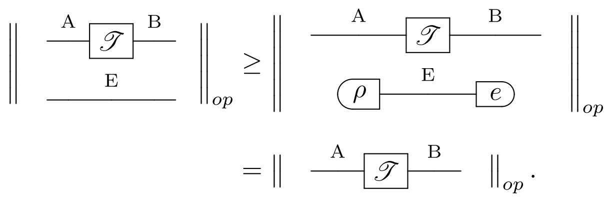
Lemma 2 (Invariance of the norm in the presence of ancillary systems). *Given a generic OPT* Θ, *for arbitrary systems* $\mathrm{A}, \mathrm{B}, \mathrm{E} \in \mathsf{Sys}(\Theta)$ *and any generalised transformation* $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$ *it holds that*

$$\|\mathscr{T}\|_{op} = \|\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}\|_{op}.$$

Proof. From the definition of operational norm [2], it holds that

$$\|\mathscr{T}\|_{op} \geqslant \|\mathscr{T} \boxtimes \mathscr{I}_{\mathrm{E}}\|_{op}.$$

On the other hand, by considering a generic deterministic state $\rho \in \mathsf{St}_1(\mathrm{E})$ and deterministic effect $e \in \mathsf{Eff}_1(\mathrm{E})$, the monotonicity property of the norm implies that



□

Also the sup norm satisfies the property stated in lemma 2 [14, corollary 2].

Lemma 3. *Let* $\mathscr{T} \in \mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$ *and* $\mathscr{G} \in \mathsf{Transf}_{\mathbb{R}(\mathrm{B}\to\mathrm{C})}$, *then [14]*

$$\|\mathscr{G}\mathscr{T}\|_{sup} \leqslant \|\mathscr{G}\|_{sup}\|\mathscr{T}\|_{sup}.$$

2.3.1. Operational completeness for instruments

As already stated Cauchy completeness of the spaces of transformations ensures that whenever a transformation can be arbitrarily approximated by operations in an OPT then it is also an admitted transformation of the theory. Here we extend this requirement to the case of instruments.

Let us start by considering the scenario of preparations. Suppose that Alice can prepare two deterministic states $\rho, \sigma \in \mathsf{St}_1(\mathrm{A})$ and that she wants to probabilistically prepare one or the other, so that the resulting mixed state is given by $p\rho + (1-p)\sigma \in \mathsf{St}_1(\mathrm{A})$, with $p \in [0,1]$. This scenario can also be expressed through the notion of preparation test $\{p\rho, (1-p)\sigma\} \in \mathsf{Prep}(\mathrm{A})$ or, equivalently, $\{\rho', \sigma'\} \in \mathsf{Prep}(\mathrm{A})$, where the non-normalized states $\rho' = p\rho$ and $\sigma' = (1-p)\sigma$ were used.

If now a second experimenter, let us call him Bob, would like to reproduce Alice's preparation procedure, he would have to approximate the two states ρ', σ' with two other states ρ'', σ'', or analogously with a new instrument $\{\rho'', \sigma''\} \in \mathsf{Prep}(\mathrm{A})$. At this point we wonder what is the error in Bob's approximation. We assume that the error is equal to the sum of the errors made in approximating the two preparation events separately:

$$\|\{\rho', \sigma'\} - \{\rho'', \sigma''\}\|_{\mathrm{gen}} = \|\rho' - \rho''\|_{\mathrm{gen}} + \|\sigma' - \sigma''\|_{\mathrm{gen}},$$

where we refer now to a generic norm $\|\cdot\|_{\mathrm{gen}}$. We observe that this error is greater than the one that would be obtained by considering the coarse-grained state corresponding to the instrument:

$$\|\rho' - \rho''\|_{\mathrm{gen}} + \|\sigma' - \sigma''\|_{\mathrm{gen}} \geqslant \|(\rho' + \sigma') - (\rho'' + \sigma'')\|_{\mathrm{gen}}.$$

This is not unexpected since via coarse-graining one disregards information about what particular event occurred within the instrument, and this removes one source of error.

We can proceed to equip the spaces of instruments of an OPT also with a topological structure. The definition will be given by referring to the generic norm $\|\cdot\|_{\mathrm{gen}}$, reminding that the cases of interest for the framework are those where the norm is equal to the operational or sup one.

Definition 5 (Norm of instruments). Let Θ be an OPT equipped with a generic norm $\|\cdot\|_{\text{gen}}$ and $\{\mathscr{T}_n\}_{n\in N} \in \mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A}\to\mathrm{B})$. The norm of $\{\mathscr{T}_n\}_{n\in N}$ is defined as

$$\left\|\{\mathscr{T}_n\}_{n\in N}\right\|_{\text{gen}} := \sum_{n\in N} \|\mathscr{T}_n\|_{\text{gen}}. \tag{7}$$

In the case of instruments, the operational norm in (7) does not carry the same operational interpretation as in the case of transformations. Even though it gives an indication as to how close two instruments are, it is not related to the optimal discrimination strategy that could be implemented. The latter would require to solve a minimax problem, whose analysis is typically hard, and amenable only under special circumstances is left for future work.

Based on the topological structure we extend the requirement of Cauchy completeness also for instruments.

3. Properties of Cauchy sequences

In this section we analyse the properties of Cauchy sequences of instruments and transformations and how they are related within an OPT in terms of a generic norm.

As a first step we have to define a Cauchy sequence of instruments $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}^n}\right\}_{n\in\mathbb{N}}$. The most formal way to see it is as a sequence of subsets $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}^n}\right\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$. One has to distinguish two cases:

1. $\exists N\in\mathbb{N}$ such that $|\mathsf{X}^n| \leqslant N$, $\forall n\in\mathbb{N}$. In this case, it is possible to treat $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}^n}\right\}_{n\in\mathbb{N}}$ as a sequence of elements in $\mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A}\to\mathrm{B})$, where everything is well-defined. This is the case we will consider in the following. Furthermore, it is possible to drop the dependence of the outcomes space from n, since, thanks to (5), it is always possible to embed every outcome space in one of cardinality N: $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}}$.
2. $\nexists N\in\mathbb{N}$ such that $|\mathsf{X}^n| \leqslant N$, $\forall n\in\mathbb{N}$. This case cannot be treated within the hypothesis assumed in the present work. The fact that two comparable instruments must have outcome spaces with equal cardinality imposes that the cardinality of the outcome spaces of the instruments in the sequence stabilizes to a fixed value. We highlight that the latter hypothesis also guarantees that the cardinality of the outcome space of the limit instrument is finite.

We can now state the completeness result for instruments and a consequent corollary on their coarse-graining.

Theorem 1. *Given a generic OPT,* $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}_{\mathbb{R}}^{(N)}(\mathrm{A}\to\mathrm{B})$ *is Cauchy with respect to the norm* $\|\cdot\|_{gen}$ *if and only if each sequence of transformations* $\{\mathscr{T}_x^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$ *is Cauchy with respect to that norm* $\forall x\in\mathsf{X}$. *Furthermore* $\lim_{n\to\infty}\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\} = \{\lim_{n\to\infty}\mathscr{T}_x^n\}_{x\in\mathsf{X}}$.

Proof. Let us start by proving the first statement. $\Longrightarrow$) Let $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}}$ be a Cauchy sequence, then $\forall\varepsilon$ there exist $\tilde{n}$ such that $\forall n,m\geqslant\tilde{n}$ it holds that

$$\left\|\{\mathscr{T}_x\}^n_{x\in\mathsf{X}} - \{\mathscr{T}_x\}^m_{x\in\mathsf{X}}\right\|_{\text{gen}} = \sum_{x\in\mathsf{X}} \|\mathscr{T}_x^n - \mathscr{T}_x^m\|_{\text{gen}} \leqslant \varepsilon$$

which implies

$$\|\mathscr{T}_x^n - \mathscr{T}_x^m\|_{\text{gen}} \leqslant \varepsilon \quad \forall x\in\mathsf{X}.$$

$\Longleftarrow$) On the other hand if the sequences of transformations are Cauchy, i.e. $\forall\varepsilon$ exists $\tilde{n}$ such that $\forall n,m\geqslant\tilde{n}$ it holds that

$$\|\mathscr{T}_x^n - \mathscr{T}_x^m\|_{\text{gen}} \leqslant \varepsilon \quad \forall x\in\mathsf{X},$$

then it immediately follows

$$\left\|\{\mathscr{T}_x\}^n_{x\in\mathsf{X}} - \{\mathscr{T}_x\}^m_{x\in\mathsf{X}}\right\|_{\text{gen}} = \sum_{x\in\mathsf{X}} \|\mathscr{T}_x^n - \mathscr{T}_x^m\|_{gen} \leqslant |\mathsf{X}|\varepsilon,$$

which, being the cardinality of X ($|\mathsf{X}|$) equal to $N<\infty$, implies that $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}}$ is Cauchy. We then conclude the proof by showing that the limit of the Cauchy sequence of the collection is given by the collection of the limits. Let indeed $\{\mathscr{T}_x\}_{x\in\mathsf{X}}$ be the limit of the Cauchy sequence $\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\}_{n\in\mathbb{N}}$—that exists by virtue of completeness of instruments. Then $\forall\varepsilon,\exists\tilde{n}$ such that $\forall n\geqslant\tilde{n}$

$$\left\|\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}-\{\mathscr{T}_x\}_{x\in\mathsf{X}}\right\|_{\mathrm{gen}}=\sum_{x\in\mathsf{X}}\|\mathscr{T}_x^n-\mathscr{T}_x\|_{\mathrm{gen}}\leqslant\varepsilon,$$

which, by the completeness requirement for transformations, implies that each sequence of events $\{\mathscr{T}_x^n\}_{n\in\mathbb{N}}$ converges to $\mathscr{T}_x$. □

Corollary 1. *Given a generic OPT and a Cauchy sequence* $\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\}_{n\in\mathbb{N}}\subset\mathsf{Instr}^{(N)}_{\mathbb{R}}(\mathrm{A}\to\mathrm{B})$ *with respect to a generic norm* $\|\cdot\|_{gen}$, *each sequence of coarse-grainings* $\{\sum_{x\in\mathsf{X}'}\mathscr{T}_x^n\}_{n\in\mathbb{N}}\subset\mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$ *is Cauchy with respect to that norm* $\forall\mathsf{X}'\subseteq\mathsf{X}$. *Furthermore* $\lim_{n\to\infty}\left(\sum_{x\in\mathsf{X}'}\mathscr{T}_x^n\right)=\sum_{x\in\mathsf{X}'}\left(\lim_{n\to\infty}\mathscr{T}_x^n\right)$.

Proof. Using theorem 1, the sequences of transformations $\{\mathscr{T}_x^n\}_{n\in\mathbb{N}}$ are Cauchy for each $x\in\mathsf{X}$. Therefore, $\forall\varepsilon$ exists $\tilde{n}$ such that $\forall n,m\geqslant\tilde{n}$ it holds that

$$\|\mathscr{T}_x^n-\mathscr{T}_x^m\|_{op}\leqslant\varepsilon\quad\forall x\in\mathsf{X},$$

then it immediately follows that

$$\left\|\sum_{x\in X'}(\mathscr{T}_x^n-\mathscr{T}_x^m)\right\|_{op}\leqslant\sum_{x\in\mathsf{X}'}\|\mathscr{T}_x^n-\mathscr{T}_x^m\|_{op}\leqslant|\mathsf{X}'|\varepsilon,$$

and being $|\mathsf{X}'|\leqslant|\mathsf{X}|=N$, this implies that $\{\sum_{x\in\mathsf{X}'}\mathscr{T}_x^n\}_{n\in\mathbb{N}}$ is Cauchy. The second statement immediately follows from the possibility of exchanging limits and summations. □

The next two lemmas consider the limit of Cauchy sequences of deterministic transformations.

Lemma 4. *In a generic OPT the limit of a Cauchy sequence, in the operational norm, of deterministic transformations is a deterministic transformation. In other words,* $\mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}$ *is closed with respect to the operational norm for any* $\mathrm{A},\mathrm{B}\in\mathsf{Sys}(\Theta)$.

Proof. The first step is to prove that the spaces of deterministic states are closed with respect to the operational norm. To show this it is sufficient to show that the limit $\rho=\lim_{n\to\infty}\rho^n$, where $\{\rho^n\}_{n\in\mathbb{N}}\subset\mathsf{St}_1(\mathrm{A})$ gives probability equal to one when evaluated on any deterministic effect $\mathrm{e}\in\mathsf{Eff}_1(\mathrm{A})$ (this is a necessary and sufficient condition for a state to be deterministic [11]). From the monotonicity of the operational norm (lemma 1) it follows that

$$\|(\mathrm{e}|\,\rho)_{\mathrm{A}}-(\mathrm{e}|\,\rho^n)_{\mathrm{A}}\|_{op}\leqslant\|\rho-\rho^n\|_{op}.$$

The latter relation then implies that

$$\lim_{n\to\infty}(\mathrm{e}|\,\rho^n)_{\mathrm{A}}=(\mathrm{e}|\,\rho)_{\mathrm{A}}\quad\forall\mathrm{e}\in\mathsf{Eff}_1(\mathrm{A}).$$

Since $(\mathrm{e}|\,\rho^n)_{\mathrm{A}}=1$ for all $n\in\mathbb{N}$, it immediately follows $(\mathrm{e}|\,\rho)_{\mathrm{A}}=1$ for all $\mathrm{e}\in\mathsf{Eff}_1(\mathrm{A})$.

To prove the result in the case of transformations, we will use the fact that a transformation $\mathscr{T}\in\mathsf{Transf}(\mathrm{A}\to\mathrm{B})$ is deterministic if and only if $(\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}})|\rho)_{\mathrm{AE}}\in\mathsf{St}_1(\mathrm{BE})$ for any $\rho\in\mathsf{St}_1(\mathrm{AE})$, where E is a generic system of the theory. Then, let $\mathscr{T}=\lim_{n\to\infty}\mathscr{T}^n$, with $\{\mathscr{T}^n\}_{n\in\mathbb{N}}\subset\mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}$. From lemmas 1 and 2, it immediately follows that $\{(\mathscr{T}^n\boxtimes\mathscr{I}_{\mathrm{E}})|\rho)_{\mathrm{AE}}\}_{n\in\mathbb{N}}\subset\mathsf{St}_1(\mathrm{BE})$, for any system E of the theory and any deterministic state $\rho\in\mathsf{St}_1(\mathrm{AE})$, is still a Cauchy sequence that converges to

$$(\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}})|\rho)_{\mathrm{AE}}=\lim_{n\to\infty}(\mathscr{T}^n\boxtimes\mathscr{I}_{\mathrm{E}})|\rho)_{\mathrm{AE}}\in\mathsf{St}_1(\mathrm{BE}).$$

In conclusion of this proof, we observe that it would have been possible to carry out the second part also referring to effects. A transformation $\mathscr{T}\in\mathsf{Transf}(\mathrm{A}\to\mathrm{B})$ is deterministic if and only if $(\mathrm{e}|_{\mathrm{BE}}(\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}})\in\mathsf{Eff}_1(\mathrm{BE})$ for any $\mathrm{e}\in\mathsf{Eff}_1(\mathrm{BE})$, where E is a generic system of the theory. □

Lemma 5. *In a generic OPT the limit of a Cauchy sequence, in the operational norm, of deterministic product transformations* $\{\mathscr{T}^n\boxtimes\mathscr{G}^n\}_{n\in\mathbb{N}}\subset\mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}\boxtimes\mathsf{Transf}_{1(\mathrm{C}\to\mathrm{D})}$ *is still a deterministic product transformation. In other words, the set of deterministic product transformations is closed with respect to the operational norm.*

Proof. By lemma 4, the sequence converges to a deterministic transformation. Therefore, it is sufficient to show that the limit is a product. The first step is to show that also the sequences $\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{\mathbb{R}(\mathrm{A}\to\mathrm{B})}$ and $\{\mathscr{G}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{1(\mathrm{C}\to\mathrm{D})}$ are Cauchy. This can be done by direct calculation:

$$\begin{aligned}
&\|\mathscr{T}^n \boxtimes \mathscr{G}^n - \mathscr{T}^m \boxtimes \mathscr{G}^m\|_{op} \\
&\quad \geqslant \left\| \mathscr{T}^n \boxtimes (\mathrm{e}|_{\mathrm{D}}\mathscr{G}^n|\rho)_{\mathrm{C}} - \mathscr{T}^m \boxtimes (\mathrm{e}|_{\mathrm{D}}\mathscr{G}^m|\rho)_{\mathrm{C}} \right\|_{op} \\
&\quad = \|\mathscr{T}^n - \mathscr{T}^m\|_{op},
\end{aligned}$$

where $\rho \in \mathsf{St}_1(\mathrm{C})$ and $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{D})$ and the inequality follows from lemma 1, and similarly for the case of $\mathscr{G}^n$. Defining $\mathscr{T} := \lim_{n\to\infty} \mathscr{T}^n$ and $\mathscr{G} = \lim_{n\to\infty} \mathscr{G}^n$, and exploiting lemma 1, one can prove that

$$\|\mathscr{T} \boxtimes \mathscr{G} - \mathscr{T}^n \boxtimes \mathscr{G}^n\|_{op} \leqslant \|\mathscr{T} - \mathscr{T}^n\|_{op} + \|\mathscr{G} - \mathscr{G}^n\|_{op},$$

which implies

$$\mathscr{T} \boxtimes \mathscr{G} = \lim_{n\to\infty} \mathscr{T}^n \boxtimes \mathscr{G}^n.$$

□

3.1. Compositional structure and coarse-graining vs Cauchy completion

By definition, an OPT is closed under sequential and parallel composition, and all the sets of instruments and transformations are closed in the topology induced by the operational norm. However, in the process of *constructing* a theory element by element, one needs to check consistency of the mathematical structures with the axioms of the framework. We then consider whether closure under sequential and parallel composition, as well as under coarse-graining, is preserved after Cauchy-completing tentative spaces of instruments and transformations with respect to the operational norm. The following results provide useful insights in this context.

For simplicity let us denote with Θ the OPT obtained from a tentative theory $\tilde{\Theta}$ through the operation of completion. Thanks to corollary 1 we know that the spaces of transformations of Θ are still closed under the operation of coarse-graining. We have now to check what happens in the case of sequential and parallel composition.

Let us start with the former. We will present the argument in the case of transformations, but it is immediately extended to the case of instruments through theorem 1. Let $\mathscr{T} \in \mathsf{Transf}(\mathrm{A}\to\mathrm{B})$ and $\mathscr{G} \in \mathsf{Transf}(\mathrm{B}\to\mathrm{C})$ be two transformations of Θ. We have to prove that $\mathscr{G} \circ \mathscr{T}$ is also a transformation of Θ. To show this, let us consider $\mathscr{T} \in \overline{\mathsf{Transf}(\mathrm{A}\to\mathrm{B})}$ as the limit of a Cauchy sequence of transformations $\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}(\mathrm{A}\to\mathrm{B})$, i.e. $\mathscr{T} = \lim_{n\to\infty} \mathscr{T}^n$, and analogously for $\mathscr{G} = \lim_{n\to\infty} \mathscr{G}^n$. We can then rewrite

$$\mathscr{G} \circ \mathscr{T} = \lim_{m\to\infty} \mathscr{G}^m \circ \lim_{n\to\infty} \mathscr{T}^n.$$

To prove the result, we have then to show that $\mathscr{G} \circ \mathscr{T}$ is the limit of a Cauchy sequence of transformations of $\tilde{\Theta}$ by showing that

$$\mathscr{G} \circ \mathscr{T} = \lim_{m\to\infty} \lim_{n\to\infty} \mathscr{G}^m \circ \mathscr{T}^n.$$

To this end, we will exploit the fact that the operational and sup norm are equivalent, which implies that

$$\begin{aligned}
\|\mathscr{G}\mathscr{T}\|_{op} \leqslant C\|\mathscr{G}\mathscr{T}\|_{sup} &\leqslant C\|\mathscr{G}\|_{sup}\|\mathscr{T}\|_{sup} \\
&\leqslant \frac{C}{c^2}\|\mathscr{G}\|_{op}\|\mathscr{T}\|_{op},
\end{aligned} \tag{8}$$

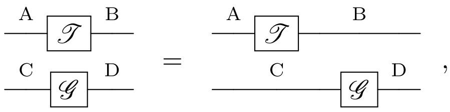
where c and C are suitable positive real constants, and the second inequality is obtained by exploiting lemma 3. Using the latter equation it then follows that:

$$\begin{aligned}
&\|\mathscr{G}^m\mathscr{T}^n - \mathscr{G}\mathscr{T}\|_{op} \\
&\quad = \|\mathscr{G}^m\mathscr{T}^n - \mathscr{G}^m\mathscr{T} + \mathscr{G}^m\mathscr{T} - \mathscr{G}\mathscr{T}\|_{op} \\
&\quad \leqslant \|\mathscr{G}^m(\mathscr{T}^n - \mathscr{T})\|_{op} + \|(\mathscr{G}^m - \mathscr{G})\mathscr{T}\|_{op} \\
&\quad \leqslant \frac{C}{c^2}\|\mathscr{G}^m\|_{op}\|\mathscr{T}^n - \mathscr{T}\|_{op} + \frac{C}{c^2}\|\mathscr{G}^m - \mathscr{G}\|_{op}\|\mathscr{T}\|_{op} \\
&\quad \leqslant \frac{C}{c^2}\left(\|\mathscr{T}^n - \mathscr{T}\|_{op} + \|\mathscr{G}^m - \mathscr{G}\|_{op}\right).
\end{aligned}$$

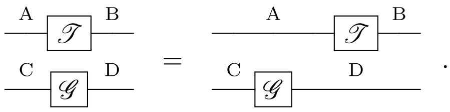
Given that both $\{\mathscr{G}^m\}_{m\in\mathbb{N}}$ and $\{\mathscr{T}^n\}_{n\in\mathbb{N}}$ are Cauchy, the latter series of inequalities implies that $\{\mathscr{G}^m\mathscr{T}^n\}_{n,m\in\mathbb{N}}$ is Cauchy and converges to $\mathscr{G}\mathscr{T}$.

In the last inequality we used the fact that the operational norm of a generic transformation $\mathscr{T}\in\mathsf{Transf}(\mathrm{A}\rightarrow\mathrm{B})$, for any couple of systems $\mathrm{A},\mathrm{B}\in\mathsf{Sys}(\Theta)$, is always bounded by one. This can be checked by direct calculation using the explicit formula for the norm provided in [2, 14].

With respect to the closure under parallel composition, the result follows considering that two operations composed in parallel can be always seen as the sequential composition of the same transformations composed with the identity. In diagrams:



or equivalently



Remembering that $\|\mathscr{T}\boxtimes\mathscr{I}_{\mathrm{E}}\|_{op}=\|\mathscr{T}\|_{op}$ for any system E of the theory (lemma 2) one can reduce to the case of sequential composition.

In summary, we proved that completing a tentative OPT including limits of Cauchy sequences of the instruments and transformations spaces returns a well defined OPT whenever the operational and sup norm are equivalent. The result is thus the following lemma.

Lemma 6. *Let $\tilde{\Theta}$ be a tentative OPT whose instruments and transformations spaces are not Cauchy complete. Then, the theory Θ obtained by Cauchy completing these spaces is a well defined OPT provided that the operational and sup norm are equivalent.*

The argument we just exposed also shows another important property that OPTs satisfy whenever the operational and sup norm are equivalent.

Lemma 7. *Let Θ be an OPT where the operational and sup norms are equivalent. Consider the two Cauchy sequences of instruments $\left\{\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}\right\}_{n\in\mathbb{N}}\subset\mathsf{Instr}(\mathrm{A}\rightarrow\mathrm{B})$ and $\left\{\{\mathscr{G}_y\}^n_{y\in\mathsf{Y}}\right\}_{n\in\mathbb{N}}\subset\mathsf{Instr}(\mathrm{B}\rightarrow\mathrm{C})$, whose limits are $\{\mathscr{T}_x\}_{x\in\mathsf{X}}=\lim_{n\rightarrow\infty}\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}$ and $\{\mathscr{G}_y\}_{y\in\mathsf{Y}}=\lim_{n\rightarrow\infty}\{\mathscr{G}_y\}^n_{y\in\mathsf{Y}}$, respectively. Then, the composite limit instrument is given by the composition of the limits*

$$\{\mathscr{G}_y\}_{y\in\mathsf{Y}}\circ\{\mathscr{T}_x\}_{x\in\mathsf{X}}=\lim_{m\rightarrow\infty}\lim_{n\rightarrow\infty}\{\mathscr{G}_y\}^m_{y\in\mathsf{Y}}\circ\{\mathscr{T}_x\}^n_{x\in\mathsf{X}}.$$

Lemma 8. *Let Θ be an OPT where the operational and sup norms are equivalent. Consider the two Cauchy sequences of transformations $\{\mathscr{T}^n\}_{n\in\mathbb{N}}\subset\mathsf{Transf}(\mathrm{A}\rightarrow\mathrm{B})$ and $\{\mathscr{G}^n\}_{n\in\mathbb{N}}\subset\mathsf{Transf}(\mathrm{B}\rightarrow\mathrm{C})$, whose limits are $\mathscr{T}=\lim_{n\rightarrow\infty}\{\mathscr{T}^n\}_{n\in\mathbb{N}}$ and $\mathscr{G}=\lim_{n\rightarrow\infty}\{\mathscr{G}^n\}_{n\in\mathbb{N}}$, respectively. Then, the composite limit transformation is given by the composition of the limits*

$$\mathscr{G}\circ\mathscr{T}=\lim_{m\rightarrow\infty}\lim_{n\rightarrow\infty}\mathscr{G}^m\circ\mathscr{T}^n.$$

These two latter results can be immediately generalised to hold for generalised instruments and transformations also, and in the case of the parallel composition.

4. Strong completeness

We prove here how to complete a theory in such a way that it possesses all the conditional operations, or more formally to a theory that is *strongly causal*.

4.1. Strongly causal OPTs

The idea is to include in the theory every instrument that can originate from conditional operations based on experimental results. Thus we refer to the most general conditional operation that can occur in a theory that is a conditional instrument.

Definition 6 (Conditional instruments). Let Θ be an OPT, $\mathsf{T}_\mathsf{X} = \{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ a test of the theory, and $\left\{\mathsf{G}^{(x)}_\mathsf{Y} = \left\{\mathscr{G}^{(x)}_y\right\}_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$ a labelled collection of instruments. We then define the following collection as a *conditional instrument*

$$\bigcup_{x\in\mathsf{X}}\left\{\mathscr{G}^{(x)}_y\circ\mathscr{T}_x\right\}_{y\in\mathsf{Y}} = \left\{\left\{\mathscr{G}^{(x)}_y\circ\mathscr{T}_x\right\}_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}}. \tag{9}$$

Remark 1. Although the outcome space of the conditional instrument may depend on x, this dependence can always be eliminated—thanks to (5)—by considering a common outcome space Y with the largest cardinality.

Remark 2. In general the object in (9) may not belong to $\mathsf{Instr}(\Theta)$, namely it may model an operation that it is not actually implementable in the theory. An example of a theory where not every conditional instrument is an actual instrument of the theory is Minimal Classical Theory (MCT) proposed in [13].

The name conditional instruments is derived from the fact that the outcome of the first test *conditions* which test the experimenter implements next[9]. We can now provide the formal definition of *strong causality*.

Definition 7 (Strongly causal OPTs). An OPT Θ is *strongly causal* if every operation of the form (9) belongs to $\mathsf{Instr}(\Theta)$, that is if every conditional instrument is an instrument of the theory.

In the literature on QT the property of strong causality is also referred to as *classical control on outcomes* and *post-processing* [52]. The name strong causality comes from the fact, that this property is stronger than the causality property usually considered within the framework:

Definition 8 (Causal OPTs). A *causal OPT* Θ is a theory where every system $\mathrm{A}\in\mathsf{Sys}(\Theta)$ admits a unique deterministic effect. In symbols, $\forall \mathrm{A}\in\mathsf{Sys}(\Theta)\ \exists!\ e\in\mathsf{Eff}_1(\mathrm{A})$ [2].

The causality condition can be proven to be equivalent to the property of *no-signalling from the future*, that is that the probability distributions of preparation-instruments does not depend on the choice of the observation-instruments at their output [2]. In causal theories the flow of information is fixed from input to output (diagrammatically, from left to right): the input–output direction can then be identified with the direction from past to future, i.e. the *arrow of time*. Furthermore, the causality condition also implies the property of *no-signalling without interaction*, namely, the fact that separated parties cannot influence each other if they do not interact. The causal structure induced by causality among parties in a network coincides with the so-called *Einstein locality* [1, 2].

4.2. Completeness with respect to strong causality

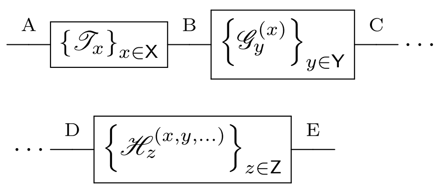
In the following theorem we show how to complete a causal theory under the strong causality assumption (remind that causality is a necessary, but not sufficient condition for strong causality):

Theorem 2 (completion under strong causality). *Let Θ be a causal OPT. If one constructs a new theory $\tilde{\Theta}$ by adding all the conditional instruments and subsequently completing the spaces of transformations and instruments endowed with the operational norm, then $\tilde{\Theta}$ is a strongly causal OPT.*

The proof will be constructive, and in particular it consists in the following extension procedure, along with a consistency check.

Procedure 1 (How to make an OPT strongly causal). The algorithm to make an OPT strongly causal consists of three different steps:

(I) Addition of all the functions obtained from conditioning a family of instruments a finite number of times. More precisely, one has to add all instruments of the form



(10)

[9] We observe that in the literature, the families of instruments depending on a classical parameter, like $\left\{\mathsf{G}^{(x)}_\mathsf{Y} = \left\{\mathscr{G}^{(x)}_y\right\}_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$, are also referred to as *multimeters* [49–51].

where $(x, y, \ldots)$ indicates the finite space of the outcomes of the instruments that precede $\left\{\mathscr{H}_z^{(x,y,\ldots)}\right\}_{z\in\mathsf{Z}}$.

(II) Closure of the new instruments and transformations spaces under sequential and parallel composition, and coarse-graining,

(III) Cauchy completion of the spaces of instruments and transformations with respect to the operational norm.

Remark 3. Coherently with the requirements stated at the beginning of section 3, the procedure described above does not lead to include in $\tilde{\Theta}$ instruments obtained by conditioning an infinite number of times. Otherwise, it could lead to the definition of sequences of instruments where the cardinality of the outcome space is not bounded.

Remark 4. The reason why in the first step of procedure 1 one is adding all instruments obtained through a finite number of conditioning steps, instead of just the instruments of the form (9)—i.e.,with just one conditioning step—as would be required by the definition of strongly causal OPTs (definition 7) is to avoid recursion in the procedure. Indeed, if one just added the instruments with one conditioning step, then the instruments obtained through conditioning on conditional instruments would be required to be added in a subsequent step, and so on. Instead, the requirement (10) already encompasses all instruments that would be obtained through the iteration procedure. To illustrate this we consider as an example conditioning of conditional instruments. The argument can then be straightforwardly generalised. Let

$$\left\{\mathscr{G}_y^{(x)} \circ \mathscr{T}_x\right\}_{(x,y)\in\mathsf{X}\times\mathsf{Y}},$$

be an instrument in $\mathsf{Instr}(\mathrm{A}\to\mathrm{E})$, where

$$\begin{aligned}\{\mathscr{T}_x\}_{x\in\mathsf{X}} &= \{\mathscr{T}_{x',x''}\}_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''}\\ &= \left\{\mathscr{T}'^{(x')}_{x''} \circ \mathscr{T}'_{x'}\right\}_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''},\end{aligned}$$

is an instrument in $\mathsf{Instr}(\mathrm{A}\to\mathrm{C})$, and

$$\begin{aligned}\left\{\left\{\mathscr{G}_y^{(x)}\right\}_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}} &= \left\{\left\{\mathscr{G}_{y',y''}\right\}^{(x',x'')}_{(y',y'')\in\mathsf{Y}'\times\mathsf{Y}''}\right\}_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''}\\ &= \left\{\left\{\mathscr{G}'^{(y')}_{y''} \circ \mathscr{G}'_{y'}\right\}^{(x',x'')}_{(y',y'')\in\mathsf{Y}'\times\mathsf{Y}''}\right\}_{(x',x'')\in\mathsf{X}'\times\mathsf{X}''}\\ &= \left\{\mathscr{G}''^{(y',x',x'')}_{y''} \circ \mathscr{G}'^{(x',x'')}_{y'}\right\}_{(x',x'',y',y'')\in\mathsf{X}'\times\mathsf{X}''\times\mathsf{Y}'\times\mathsf{Y}''},\end{aligned}$$

is an instrument in $\mathsf{Instr}(\mathrm{C}\to\mathrm{E})$. By composing the two instruments one obtains

$$\left\{\mathscr{G}''^{(y',x',x'')}_{y''} \circ \mathscr{G}'^{(x',x'')}_{y'} \circ \mathscr{T}'^{(x')}_{x''} \circ \mathscr{T}'_{x'}\right\}_{z\in\mathsf{Z}},$$

where $z\in\mathsf{Z}$ was used in place of $(x', x'', y', y'') \in \mathsf{X}' \times \mathsf{X}'' \times Y' \times Y''$, which is exactly of the form (10).

We are now in a position to prove theorem 2.

Proof. One has to show that $\tilde{\Theta}$, obtained through procedure 1, is a well defined OPT. This amounts to showing that after the operation of Cauchy completion all the conditional instruments are actual instruments of the theory and that this new collection of instruments is closed under sequential and parallel composition, and coarse-graining [14, 42]. Closure under composition and coarse-graining is guaranteed by lemmas 7, 8 and corollary 1. Therefore, the first requirement is the only one that needs to be checked explicitly. We then show that, after adding all the instruments obtained through a finite number of conditioning steps to Θ and subsequently completing with respect of the operational norm, for every instrument $\mathsf{T}_\mathsf{X} \equiv \{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ and for every family $\mathsf{G}_\mathsf{Y}^{(x)} \equiv \left\{\mathscr{G}_y^{(x)}\right\}_{y\in\mathsf{Y}} \subset \mathsf{Instr}(\mathrm{B}\to\mathrm{C})$, for any system $\mathrm{A}, \mathrm{B}, \mathrm{C} \in \mathsf{Sys}(\tilde{\Theta})$, the conditional set of operations

$$\left\{\left\{\mathscr{G}_y^{(x)} \circ \mathscr{T}_x\right\}_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}}$$

is a test of the theory.

In this case, it is sufficient to study what happens for instruments with a single conditional step since one is only interested in proving that $\tilde{\Theta}$ satisfies definition 7. Let us suppose that there exists a particular family of tests for which this does not hold. Since we already added to $\tilde{\Theta}$ all the conditional instruments that it is possible to construct using the tests of Θ, the only cases we have to consider are the ones where at least one between T_X or $\left\{\mathsf{G}_\mathsf{Y}^{(x)}\right\}_{x\in\mathsf{X}}$ is an instrument obtained as a limit from the procedure of Cauchy completion. Given that an instrument can always be seen as the limit of a constant sequence, we will treat both T_X and $\left\{\mathsf{G}_\mathsf{Y}^{(x)}\right\}_{x\in\mathsf{X}}$ as limits. Thanks to theorem 1 along with lemmas 7 and 8 we have that

$$\begin{aligned}\mathscr{G}_y^{(x)} \circ \mathscr{T}_x &= \left(\lim_{m\to\infty} \mathscr{G}_y^{(x)^m}\right) \circ \left(\lim_{n\to\infty} \mathscr{T}_x^n\right) \\ &= \lim_{m\to\infty} \lim_{n\to\infty} \mathscr{G}_y^{(x)^m} \circ \mathscr{T}_x^n.\end{aligned}$$

Hence, $\mathscr{G}_y^{(x)} \circ \mathscr{T}_x$ is the limit of a sequence of transformations of $\tilde{\Theta}$, and by the requirement of Cauchy completeness of the spaces of transformations one obtains the result, i.e.

$$\left\{\left\{\mathscr{G}_y^{(x)} \circ \mathscr{T}_x\right\}_{y\in\mathsf{Y}}\right\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{C})$$

is an actual instrument of $\tilde{\Theta}$.

□

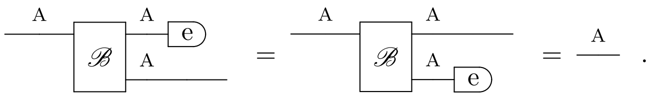
5. Broadcasting, compatibility and disturbance

In this section we introduce the notions of *broadcasting*, *compatibility of observations*, and *irreversibility of measurement disturbance*.

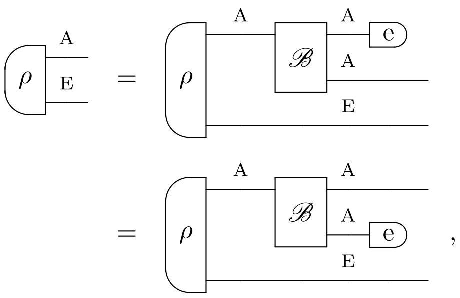
5.1. Broadcasting

Let us start from the definition of broadcasting.

Definition 9 (Broadcasting transformation). Considering a causal OPT, a transformation $\mathscr{B} \in \mathsf{Transf}_{1(\mathrm{A}\to\mathrm{AA})}$ is *broadcasting* if:

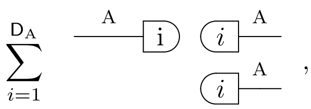

(11)

Notice that, following the above definition, a transformation $\mathscr{B} \in \mathsf{Transf}_{1(\mathrm{A}\to\mathrm{AA})}$ is broadcasting if and only if



for every system E of the theory and for every state $\rho \in \mathsf{St}(\mathrm{AE})$.

While QT does not allow for a broadcasting transformation, CT does [5, 19–26, 53] and the classical broadcasting map is given by:


(12)

where $i \in \mathsf{PurSt}(\mathrm{A})$ are the pure states of system A and $\mathrm{i} \in \mathsf{Eff}(\mathrm{A})$ are the measurements that jointly perfectly discriminate the pure states[10].

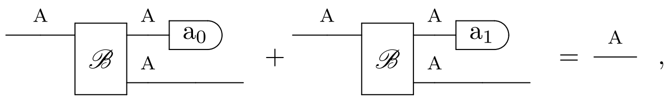
[10] A set of states $\{\rho_n\}_{n\in N}$ is *jointly perfectly discriminable* if there exists a test $\{\mathrm{a}_x\}_{x\in N} \in \mathsf{Obs}(\mathrm{A})$ such that $(\mathrm{a}_x|\,\rho_x) = 1$ for all $x \in N$.

We will say that a theory satisfies the property of *broadcasting* if every system of the theory admits a broadcasting transformation. Otherwise, we will say that the theory satisfies the *no-broadcasting* theorem.

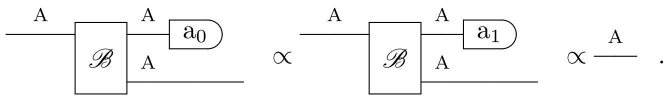
We now prove a sufficient condition for a system of an OPT to satisfy no-broadcasting.

Lemma 9. *Consider a causal OPT and a system* A *of the theory of dimension greater than 1,* $\mathsf{D}_\mathrm{A} \geqslant 2$. *If* $\mathscr{I}_\mathrm{A}$ *is atomic, then the system* A *does not admit a broadcasting channel* $\mathscr{B} \in \mathsf{Transf}_1(\mathrm{A}\to\mathrm{AA})$.

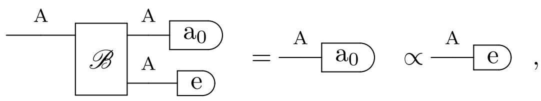
Proof. Suppose by contradiction that there exists a system A, $\mathsf{D}_\mathrm{A} \geqslant 2$, such that $\mathscr{I}_\mathrm{A}$ is atomic and there exists a broadcasting channel $\mathscr{B} \in \mathsf{Transf}_{1(\mathrm{A}\to\mathrm{AA})}$. Consider now a non-trivial decomposition $\{\mathrm{a}_0, \mathrm{a}_1\} \neq \{p_0\mathrm{e}, p_1\mathrm{e}\} \in \mathsf{Obs}(\mathrm{A})$ of the deterministic effect $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{A})$. An observation-test of this kind always exists, since otherwise the system would have dimension equal to 1. By (11), it follows that:



which, by the atomicity of the identity transformation, implies:



Exploiting again (11) one obtains that:



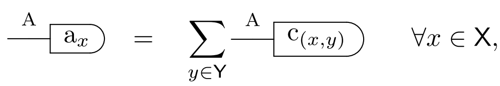
and analogously for a_1, which contradicts the hypothesis $\mathsf{D}_\mathrm{A} \geqslant 2$. □

5.2. Compatibility and irreversibility

Two operations are said to be *compatible* if operating one does not preclude the possibility of performing the other on the same system. For example, position and momentum measurements are compatible in CT as they can be performed simultaneously, but not in QT. In the following, we use 'compatibility' referring to observation-tests, and we define theories with full compatibility as those theories where all observation-tests are pairwise compatible. On the contrary, we will define *irreversible disturbance* the existence of some operation that prevents from making another one on the same system. In this sense the first operation causes an unavoidable disturbance.

Let us look at observation-instruments and identify the class of theories where all observations are compatible, as it happens for example in CT:

Definition 10 (OPTs with full-compatibility of the observation-instruments). A causal OPT Θ is said to satisfy *full-compatibility of the observation-instruments* if every pair of observation-instruments $\{\mathrm{a}_x\}_{x\in\mathsf{X}}$, $\{\mathrm{b}_y\}_{y\in\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$ of the theory, for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, are *compatible*, namely there exists a third test $\{\mathrm{c}_{(x,y)}\}_{(x,y)\in\mathsf{X}\times\mathsf{Y}} \in \mathsf{Obs}(\mathrm{A})$ such that [13, 54]

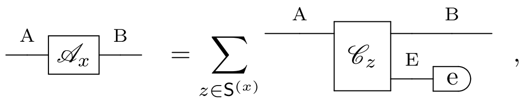


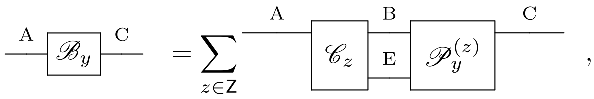
This definition can be seen as a special instance of the more complex definition involving arbitrary instruments, however in the above form it is a straightforward extension to OPTs of the definition introduced in the quantum literature [16, 17, 55]

In the following, the shorter nomenclature *compatibility* will be used in place of *full-compatibility of the observation-instruments*.

The notion of irreversibility, on the other hand, describes processes that cause an irreducible disturbance on the systems on which they act. This is based on the notion of exclusion between instruments.

Definition 11 (Does not exclude). Let Θ be a causal OPT, we say that an instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ *does not exclude* another instrument $\{\mathscr{G}_y\}_{y\in\mathsf{Y}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{C})$ if there exists a test $\{\mathscr{C}_z\}_{z\in\mathsf{Z}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{BE})$ and a post-processing, i.e. a family of instruments $\left\{\left\{\mathscr{P}_y^{(z)}\right\}_{y\in\mathsf{Y}}\right\}_{z\in\mathsf{Z}} \subset \mathsf{Instr}(\mathrm{BE}\to\mathrm{C})$ such that


(13)


(14)

where $\left\{\mathsf{S}^{(x)}\right\}_{x\in\mathsf{X}}$ is a suitable partition of X [54]. On the other hand, if the above condition fails, we say that the instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}}$ *excludes* $\{\mathscr{G}_y\}_{y\in\mathsf{Y}}$.

A theory with irreversibility is defined as follows.

Definition 12 (OPTs with irreversibility). A causal OPT Θ is said to have *irreversibility of measurement disturbance* (or in short *irreversible disturbance*) if it admits an *intrinsically irreversible* instrument, i.e. a test that excludes some other test of the theory.

Notice that the kind of disturbance introduced by an intrinsically irreversible instrument is the 'strictest' possible: if an instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}}$ excludes $\{\mathscr{G}_y\}_{y\in\mathsf{Y}}$, one cannot obtain $\{\mathscr{G}_y\}_{y\in\mathsf{Y}}$ even using any conceivable resources, namely any kind of post-processing involving arbitrary ancillary systems. An example of theory with irreversibility is QT, where almost all quantum instruments are intrinsically irreversible. We remark, however, that since every quantum channel admits a unitary dilation involving ancillary systems, all quantum channels are not intrinsically irreversible [13].

On the contrary, CT does not have irreversibility. This implies that whatever operation in CT can be implemented in such a way to reverse it and restore the initial state of the system.

Two sufficient conditions for irreversibility have been proved in [13]. The first is given by the following lemma:

Lemma 10. *An instrument is intrinsically irreversible if and only if it excludes the identity [13].*

Another condition for irreversibility is the atomicity of the identity transformation[11]:

Theorem 3. *Let Θ be a causal OPT with a system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ such that $\mathsf{D}_\mathrm{A} \geqslant 2$, and let its identity transformation $\mathscr{I}_\mathrm{A}$ be atomic. Then, there exists an instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$, for some system $\mathrm{B} \in \mathsf{Sys}(\Theta)$ that is intrinsically irreversible. Hence, the theory has irreversibility [18].*

The proof of the above theorem is also reported in appendix A for completeness.

A particular class of theories that satisfy the hypothesis of theorem 3 is that of OPTs where the identity transformation is atomic for every system of the theory, which is equivalent to n*o-information without disturbance (NIWD)* [18]. The last property means that in order to extract non-trivial information from a system one must necessarily disturb it. QT is an example of a theory with NIWD, which in turn underlies the possibility of devising information-theoretically secure cryptographic protocols, since any intervention from an eavesdropper would be detectable by the other parties [56–64].

Regarding the relation between the notions of incompatibility of measurements and of processes (named irreversibility) it has been proven in [13] that while

$$\text{incompatibility} \Rightarrow \text{irreversibility},$$

on the other hand

$$\text{incompatibility} \not\Leftarrow \text{irreversibility}.$$

A counterexample to the second implication is MCT [13], which has full-compatibility of the observation-instruments, yet it admits intrinsically irreversible tests.

[11] We remark that the identity transformation is atomic if every instrument whose full coarse-graining is the identity is of the form $\{p_x\mathscr{I}\}_{x\in\mathsf{X}}$, with $\{p_x\}_{x\in\mathsf{X}}$ a probability distribution (definition 1).

6. Minimal operational probabilistic theories

We will now proceed to introduce and characterise the class of minimal operational probabilistic theories (MOPTs). The idea is to explore the diametrically opposed situation as to what is typically considered in the probabilistic theories scenario. Indeed the set of operations in a theory is often taken as large as possible, so for example given the set of states, all maps sending legitimate states to legitimate states are included as maps of the theory. However, no bottlenecks to consistent theories with restricted set of instruments and transformations occur in principle. On the contrary, here we look at the 'smallest conceivable set' of operations, in the sense that the removal of any of the instruments would no longer lead to a legitimate OPT.

The first point we need to analyse concerns the structure of composite systems. In a general theory, systems can, in principle, be obtained in multiple ways by composing other systems, except for those that do not admit any decomposition, which we term *elementary*.

More formally, let us start from the following definition

Definition 13 (Elementary systems and minimal decomposition). We say that a system A is *elementary* if $\mathrm{A} = \mathrm{BC}$ implies $\mathrm{B} = \mathrm{I}$ or $\mathrm{C} = \mathrm{I}$. Given a system S, we say that $\mathrm{A} = \mathrm{A}_1\mathrm{A}_2\ldots\mathrm{A}_k$ is a *minimal decomposition* of S in elementary systems if A_i is elementary and non trivial for every i.

In the following, we require the decomposition in elementary systems to be unique up to operational equivalence of the single elementary systems.

Definition 14 (Unique decomposition OPTs). We say that an OPT has *unique decomposition* if for every system S, given two minimal decompositions in elementary systems $\mathrm{S} = \mathrm{A}_1\mathrm{A}_2\ldots\mathrm{A}_k$ and $\mathrm{S} = \mathrm{B}_1\mathrm{B}_2\ldots\mathrm{B}_l$ implies $k = l$ and $\mathrm{B}_i = \mathrm{A}_i, \forall i = 1,\ldots,k$.

Finally, we can define minimal theories in the following way:

Definition 15 (Minimal Operational Probabilistic Theory). We define a *Minimal Operational Probabilistic Theory (MOPT)* as an OPT with unique decomposition where the only allowed tests are the ones obtainable by composing the elements of

$$\left\{\mathsf{I}_{\star}^{\mathrm{A}\to\mathrm{A}}\,,\ \mathsf{S}_{\star}^{\mathrm{AB}\to\mathrm{BA}}\,,\ \left(S^{-1}\right)_{\star}^{\mathrm{BA}\to\mathrm{AB}}\,,\ \rho_{\mathsf{X}}\,,\ \mathsf{a}_{\mathsf{X}}\right\}, \tag{15}$$

where $\rho_{\mathsf{X}} \in \mathsf{Prep}(\mathrm{A})$ and $\mathsf{a}_{\mathsf{X}} \in \mathsf{Obs}(\mathrm{A})$ are all the possible preparation- and observation-tests of the theory, and the Cauchy completion of the aforementioned set. Thus the only allowed events are those obtainable by sequential and parallel composition of the elements of

$$\left\{\mathscr{I}_{\mathrm{A}}\,,\ \mathscr{S}_{\mathrm{A,B}}\,,\ \mathscr{S}_{\mathrm{A,B}}^{-1}\,,\ \rho\,,\ \mathsf{a}\right\}, \tag{16}$$

for every $\mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta)$, $\rho \in \mathsf{Event}(\mathrm{I}\to\mathrm{A})$ and $\mathsf{a} \in \mathsf{Event}(\mathrm{A}\to\mathrm{I})$, and the Cauchy completion of spaces of events of this type that belong to a test of the theory.

Remark 5. The above class is well-defined. One should check that the spaces of instruments and transformations are closed under the operations of sequential and parallel composition, and under the operation of coarse-graining. While the closure under coarse-graining is guaranteed by corollary 1, the closure under the compositional structure is guaranteed by lemmas 7 and 8 under the here assumed hypothesis that the operational and sup norm are equivalent.

Similar restrictions on a theory, in analogy to the ones we consider for MOPTs, have been proposed in [50, 65]. Selby *et al* [50] studies *accessible GPT fragments* which are designed to describe scenarios where states and effects are limited to just those accessible in a particular experimental setting. Even though the notion of accessible GPT fragments is close to the notion of MOPTs, there is an important difference. MOPTs, despite the restrictions, are still fully-fledged theories, while accessible GPT fragments are not in general GPT themselves. One aspect in which these two definitions differ is that of the state and effect spaces of a MOPT must be separating from each other, while this may not be true in the case of accessible GPT fragments. In [65] a class of operational theories is introduced, called *non-free*, where restrictions on operations are not followed by closure with respect to parallel and sequential composition.

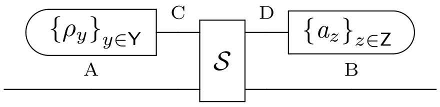
6.1. Characterization of instruments and transformations

As noted in remark 5, any MOPT is formally well-defined, although the transformations and instruments are provided only implicitly. In this section, we prove structural theorems that specify the circuital realization of the theory's instruments and transformations. Some of these results were previously presented in the supplementary material of [13], but they are derived here in a more general framework. The proofs will be included in the appendix.

Let us start with the complete characterization of the instruments and transformations that are obtainable by composing sequentially and in parallel the elements of (15) and (16), respectively, postponing the analysis of instruments and transformations obtained as limits of Cauchy sequences.

Then one has the following structural results.

Lemma 11. *In every MOPT any instrument* $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ *obtained as parallel and sequential composition of the elements of (15) is of the form:*



, (17)

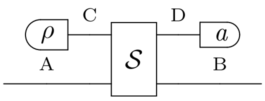
where $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AC}\to\mathrm{DB})$ *is a suitable braid transformation,* $\{\rho_y\}_{y\in\mathsf{Y}} \in \mathsf{Prep}(\mathrm{C})$, $\{a_z\}_{z\in\mathsf{Z}} \in \mathsf{Obs}(\mathrm{A})$, *the outcome space* $\mathsf{X} = \mathsf{Y}\times\mathsf{Z}$, *and* A, B, C, D $\in \mathsf{Sys}(\Theta)$ *may also be equal to the trivial system* [13].

Formally braid transformations are defined in the following way:

Definition 16 (Set of braid transformations). The *set of braid transformations*, whose representatives will be indicated with $\mathcal{S}$, is defined as the ensemble of transformations which are obtained by parallel and sequential composition of the braiding and identity transformations.

The proof of the above theorem can be found in appendix C and from it the analogous result for transformations immediately follows.

Corollary 2. *In every MOPT any transformation* $\mathscr{T} \in \mathsf{Transf}(\mathrm{A}\to\mathrm{B})$ *obtained as parallel and sequential composition of the elements of (16) is of the form:*



, (18)

where $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AC}\to\mathrm{DB})$ *is a suitable braid transformation,* $\rho \in \mathsf{St}(\mathrm{C})$, $a \in \mathsf{Eff}(\mathrm{A})$, *and* A, B, C, D $\in \mathsf{Sys}(\Theta)$ *may also be equal to the trivial system [13].*

An important property that will be used throughout the discussion is that (18) is invariant under parallel and sequential composition [13].

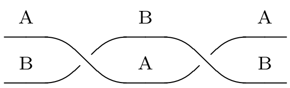
In the case of symmetric MOPTs, braid transformations become *permutations*, which allows us to further specialise the characterisation of the form of instruments and transformations, thanks to the following results.

First, let us observe that

Lemma 12. *Consider a symmetric OPT with unique decomposition and let* $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{A}\to\mathrm{B})$ *that permutes the systems as* $\mathrm{B}_j = \mathrm{A}_{\sigma(i)}$, *where* $\mathrm{A} = \mathrm{A}_1\ldots\mathrm{A}_n$ *and* $\mathrm{B} = \mathrm{B}_1\ldots\mathrm{B}_n$ *are the unique decompositions of* A *and* B, *respectively. Then the action of* $\mathcal{S}$ *is completely characterised by the permutation* σ.

Proof. The proof is a straightforward application of the coherence theorem for symmetric monoidal categories, to which symmetric OPTs belong. The theorem guarantees that the operations obtained through the composition of swap are completely defined by how they permute the order of their input objects [46, 66]. □

Remark 6. In the case of non-symmetric OPTs for a generic braid transformations the reordering of the subsystems does not completely characterise the action of the transformation. Consider, for example, the identity $\mathscr{I}_{\mathrm{AB}}$ and the following transformation:



.

Despite reordering the systems in the same way, the two transformations are generally different.

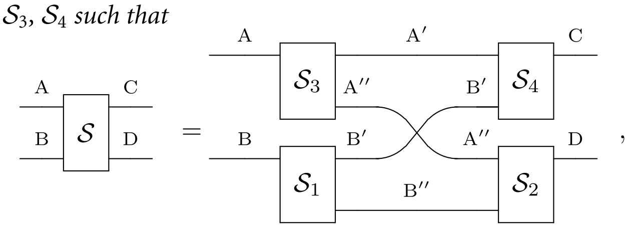
Remark 7. Also in the case in which uniqueness of decomposition does not hold, it is impossible to completely characterise a permutation by how it permutes elementary systems. Let S be a system admitting two different decompositions in elementary systems AB = CD such that A ≠ C and A ≠ D, and analogously for B. Consider then the following permutation

(19)

where $\mathscr{I}_{AB\to CD}$ is the identity test for AB = CD. Given that no relation is known between the systems BA and DC, other than the fact that they are operationally equivalent, it is not possible to state that this permutation can be completely characterised by how it permutes its input systems.

From lemma 12 the next result immediately follows.

Lemma 13 (Permutations on bipartite systems). *In every symmetric OPT with unique decomposition, for any permutation acting on a bipartite system* AB *there exist suitable systems* A', B', A'', B'', *and reversible transformations* $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$ *such that*

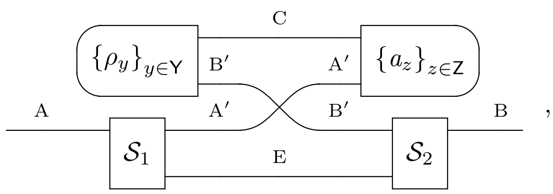


(20)

where A, B *are generic systems of the theory and* C, D *are systems such that* CD *has the same decomposition in elementary systems as* AB. *In general, any of* A, B, C, D *can be the trivial system, and the same holds also for* A', A'', B', B''.

We have now all the elements to prove the following characterisation theorems for instruments and transformations in symmetric MOPTs.

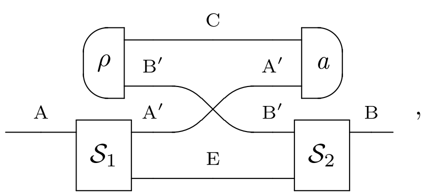
Theorem 4. *In every symmetric MOPT any instrument* $\{\mathscr{T}_x\}_{x\in X} \in \mathsf{Instr}(A \to B)$ *obtained as parallel and sequential composition of the elements of (15) is of the form:*



(21)

where $\mathcal{S}_1, \mathcal{S}_2 \in \mathsf{RevTransf}(\Theta)$ *are suitable permutations,* $\{\rho_y\}_{y\in \mathsf{Y}s} \in \mathsf{Prep}(CB')$, $\{a_z\}_{z\in Z} \in \mathsf{Obs}(CA')$, *the outcome space* $X = Y \times Z$, *and* $A, B, A', B', C, E \in \mathsf{Sys}(\Theta)$ *may also be equal to the trivial system [13].*

Corollary 3. *In every symmetric MOPT any transformation* $\mathscr{T} \in \mathsf{Transf}(A \to B)$ *obtained as parallel and sequential composition of the elements of (16) is of the form:*



(22)

where $\mathcal{S}_1, \mathcal{S}_2 \in \mathsf{RevTransf}(\Theta)$ *are suitable permutations,* $\rho \in \mathsf{St}(CB')$, $a \in \mathsf{Eff}(CA')$, *and* $A, B, A', B', C, E \in \mathsf{Sys}(\Theta)$ *may also be equal to the trivial system [13].*

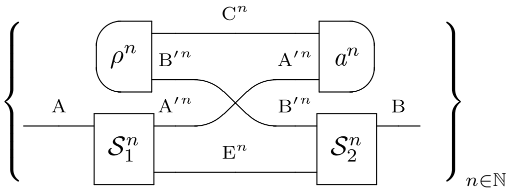
The diagrams in (21) and (22), are colloquially referred to as *jellyfish* instruments and transformations, respectively.

6.1.1. Limits of Cauchy sequences for deterministic instruments

The explicit form of instruments and transformations given by limits of Cauchy sequences of (15) and (16), which have to be included as physical objects by the completeness requirement, remains an open question. However, in the special case of symmetric causal theories, a structure theorem can be proved for Cauchy sequences of *deterministic* transformations.

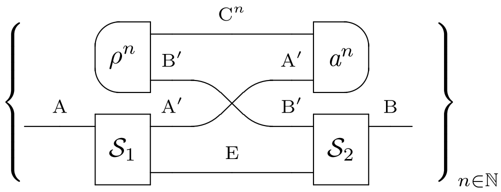
A first tool in this direction is the following lemma that holds for all symmetric MOPT (not necessarily causal).

Lemma 14. *In a symmetric MOPT any Cauchy sequence of transformations obtained as parallel and sequential composition of the elements in* (16)*,*



, (23)

admits a subsequence where the systems E^n, A'^n, B'^n *and the where the permutations* $\mathcal{S}_1^n$, $\mathcal{S}_2^n$ *are fixed:*

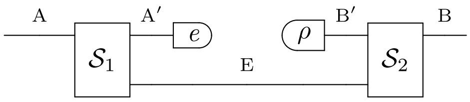


.

Remark 8. Thanks to theorem 1, the result of lemma 14 holds also in the case of instruments.

Then we have a more detailed version of corollary 3 in the case of deterministic transformations of a causal theory (not necessarily symmetric)

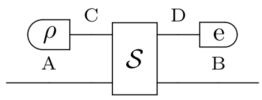
Lemma 15. *In a causal MOPT (also non-symmetric) every deterministic transformation obtained as composition of the elements in* (16) *is of the form* [13]



. (24)

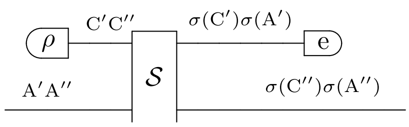
Proof. To prove the result it is sufficient to use (2), lemma 12, and the uniqueness of the deterministic effect.

The most general deterministic transformation of an MOPT is given by



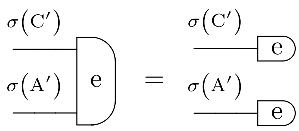
,

where $\rho \in \mathsf{St}_1(\mathrm{C})$ and $\mathrm{e} \in \mathsf{Eff}_1(\mathrm{D})$. This immediately follows from lemma 11. Given that we are considering also the case of non-symmetric MOPTs, lemma 12 cannot be exploited in this case. However, even though a generic braid transformation cannot be completely characterised by how it permutes its input wires (remark 6), its action is still of permuting them in some way. Consequently, one has



,

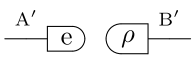
where we used $\sigma(E) = \sigma(\mathrm{E}_1 \ldots \mathrm{E}_n)$ as a shorthand notation for $\mathrm{E}_{\sigma(1)} \ldots \mathrm{E}_{\sigma(n)}$. Using the uniqueness of the deterministic effect, which implies



,

and the naturality property of the braiding (2), one obtains the desired result. □

The transformation



(25)

featuring in (24) is usually referred to as *erase and prepare* since, whatever the input is, it will discard it and prepare the state ρ.

Finally, the key information is that for symmetric causal MOPTs, the form (24) remains valid even when considering the limits of sequences of deterministic transformations:

Theorem 5. *In a causal symmetric MOPT the limits of Cauchy sequences of deterministic transformations are still of the form (24) [13].*

As for corollary 3, also the last result cannot be proven in the case of non-symmetric theories due to the fact that the reordering of the subsystems does not completely characterise the braid transformations.

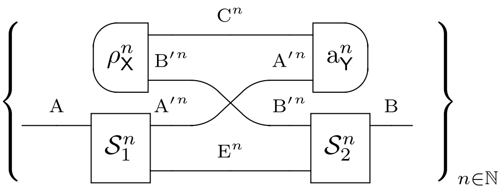
6.2. Atomicity of the identity in minimal theories

The full characterization of limits of deterministic transformations after completion in theorem 5 is at the core of one of the main results of this work, that is the atomicity of the identity transformation. The last one in turn has a number of relevant consequence, most notably the NIWD property (actually the atomicity of the identity has been proved equivalent to NIWD [18]).

The relation between minimality and atomicity of the identity is proved in the following theorem:

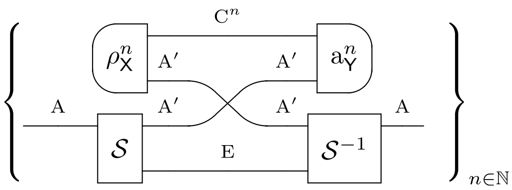
Theorem 6. *In every causal symmetric MOPT Θ the identity transformation $\mathscr{I}_{\mathrm{A}}$ is atomic for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ [13].*

Proof. To prove this result we have to show that the full coarse-graining of the limit of any Cauchy sequence of instruments of the theory that converges to the identity has to consist of transformations proportional to the identity itself. Let us start by considering a Cauchy sequence of generic instruments of the theory, which are of the form



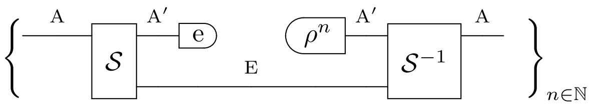
, (26)

by theorem 4. We now consider only instruments with the same input and output system since we are interested in sequences whose coarse-graining converges to the identity transformation. Lemma 14 guarantees that it is always possible to find a subsequence with the permutations and part of the systems fixed



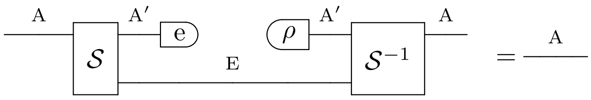
.

Consequently the sequence of full coarse-grainings is of the form



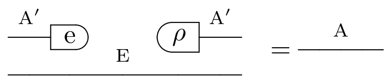
.

In order to be equal to the identity, the limit must be



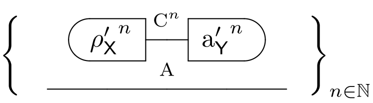
,

or equivalently



,

where the limit of the sequence is deterministic and can be obtained exploiting theorem 5, defining $\rho = \lim_{n\to\infty} \rho^n$. The last identity can hold only if $\mathrm{A}' = \mathrm{I}$ and $\mathrm{E} = \mathrm{A}$, which means that a sequence of instruments whose coarse-graining converges to the identity transformation is of the form



, (27)

obtained by substituting $A' = I$ and $E = A$ into (26). The last equation implies that every admissible decomposition of the identity is trivial, i.e.,made of transformations proportional to the identity itself. It follows that the identity is atomic. □

6.3. Properties of minimal theories

We show here the main informational consequences following the structure of minimal theories.

A first interesting fact regarding OPTs is that whenever the identity transformation is atomic for a certain system A, then any reversible transformation having that system as input or output must also be atomic. This implies that causal symmetric MOPTs do not admit reversible transformations different from permutations.

Lemma 16. *In any OPT whenever the identity transformation is atomic for a certain system* A*, then any reversible transformation having that system as input or output must also be atomic [18].*

Lemma 17. *In any causal symmetric MOPT every reversible transformation is atomic and the set of reversible transformations coincides with that of permutations.*

Proof. The first part of the corollary derives directly from theorem 6 and lemma 16. We now prove that the only reversible transformations are permutations. In the minimal setting, the only way in which a new reversible transformation, different from a permutation, can be included in the theory is as limit of some sequence of transformations. Moreover, since the set of deterministic transformations is Cauchy complete by lemma 4, and $\mathscr{R}$ is deterministic, the sequence can be taken in the set of deterministic transformations $\{\mathscr{T}^n\}_{n\in\mathbb{N}} \subset \mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}$:

$$\mathscr{R} = \lim_{n\to\infty} \mathscr{T}^n \in \mathsf{RevTransf}(\mathrm{A}\to\mathrm{B}),$$

Now consider the sequence $\{\mathscr{G}^n\}_{n\in\mathbb{N}} := \left\{\mathscr{R}^{-1}\mathscr{T}^n\right\}_{n\in\mathbb{N}}$, which is still Cauchy and converges to an instrument whose coarse-graining is $\mathscr{I}_\mathrm{A}$ as a consequence of (lemma 1) (or, equivalently, lemma 3 in the case of the sup norm). By the same argument used in the proof of theorem 6, it follows that any sequence of instruments whose coarse-graining converges to the identity consists of transformations that are proportional to the identity itself. In this specific case, where the sequence is composed of deterministic maps, all transformations must coincide with the identity, implying that the sequence is constant. Therefore, $\mathscr{I}_\mathrm{A} = \mathscr{R}^{-1} \circ \mathscr{T}^n$ for all $n \in \mathbb{N}$, which further implies that $\mathscr{R} = \mathscr{T}^n$ for any $n \in \mathbb{N}$. Since $\{\mathscr{T}^n\}_{n\in\mathbb{N}}$ is a sequence of transformations included in (16), we have thus proven that Cauchy completing a causal symmetric MOPT does not add any reversible transformation. Therefore, the set of reversible transformations coincides with the set of permutations. □

In [18] it has been shown that any theory where the identity transformation is atomic for every system satisfies the property of NIWD, namely any process that provides some non-trivial information on a system must perturb the system. Therefore, by theorem 6 we have that any minimal theory exhibit NIWD:

Corollary 4. *Every causal symmetric MOPT satisfies the property of NIWD.*

From lemma 17 and theorem 3 it immediately follows the irreversibility of minimal theories:

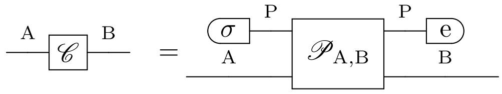
Corollary 5. *Every non-trivial causal symmetric MOPT has irreversibility*[12].

Furthermore, in every minimal theory there cannot exist a broadcasting channel, as it follows from lemma 9.

Corollary 6. *Every causal symmetric MOPT does not admit a broadcasting channel.*

In conclusion of this section, we observe that no MOPTs can satisfy the programming theorem.

Definition 17 (Universal simulator). Consider a causal OPT and a pair of systems A and B of the theory. We define a *universal simulator for* $\mathsf{Transf}(\mathrm{A}\to\mathrm{B})$, a deterministic transformation $\mathscr{P}_{\mathrm{A,B}} \in \mathsf{Transf}_{1(\mathrm{PA}\to\mathrm{PB})}$ for some suitable system P, such that for every deterministic transformation $\mathscr{C} \in \mathsf{Transf}_{1(\mathrm{A}\to\mathrm{B})}$, there exists a program state $\sigma \in \mathsf{St}_1(\mathrm{P})$ such that [1]:



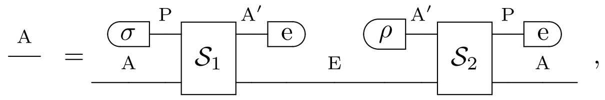
.

Definition 18 (No-programming). We say that an OPT has *no-programming* if some pair of systems A and B does not have an universal simulator.

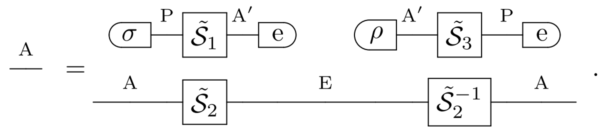
[12] We term *non-trivial* OPTs those theories that admit systems of dimension greater than 1.

Corollary 7. *Every causal symmetric MOPT has no-programming.*

Proof. By contradiction, suppose that a universal simulator for system A exists, and let us use it to program the identity. If we decompose the channel $\mathscr{P}_{\mathrm{A,B}}$ using (24), then a program $\sigma \in \mathsf{St}_1(\mathrm{P})$ for the identity must satisfy



for some suitable states and permutations. However, by the same argument used in the proof of theorem 6, the latter equality can be satisfied only if $\mathcal{S}_1 = \tilde{\mathcal{S}}_1 \boxtimes \tilde{\mathcal{S}}_2$ and $\mathcal{S}_2 = \tilde{\mathcal{S}}_3 \boxtimes \tilde{\mathcal{S}}_2^{-1}$ the following holds:



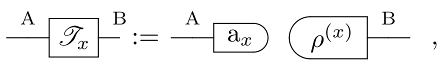
This shows that the identity can be programmed, then it is impossible to program any other transformation, thus forbidding the existence of a universal simulator. □

7. Minimal theories with strong causality

The results presented in the previous section pertain to operational theories that do not satisfy strong causality, i.e. the very natural possibility of performing conditional experiments—a feature that is reasonably expected in a theory of physical systems. This raises the question we aim to address in this section: do the characteristics of minimal theories derived so far remain valid when strong causality is assumed? To explore this, we analyse the properties of what we term minimal strongly causal operational probabilistic theories (MSOPTs), which can be understood as generic MOPTs completed to satisfy strong causality through the procedure outlined in procedure 1.

Definition 19 (Minimal Strongly Causal Operational Probabilistic Theory). We define as *Minimal Strongly Causal Operational Probabilistic Theory (MSOPT)* an OPT with unique decomposition where the allowed tests are those of the MOPT with the same systems, plus all conditional tests thereof, and finally Cauchy completed.

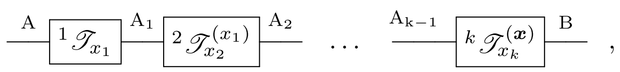
We remark that completion with respect to conditioning completely sets apart MSOPTs form MOPTs, with the set of instruments of a given minimal theory strictly contained in its completion under strong causality. For example, while the single transformations of the form



where A and B are generic systems of the theory, $\{\mathrm{a}_x\}_{x\in\mathsf{X}} \in \mathsf{Obs}(\mathrm{A})$ is a generic observation-test, and $\{\rho^{(x)}\}_{x\in\mathsf{X}} \subset \mathsf{St}_1(\mathrm{B})$ is a collection of generic deterministic states of the theory, are legitimate in MOPTs, their collection $\{\mathscr{T}_x\}_{x\in\mathsf{X}}$ is an instrument only in MSOPTs.

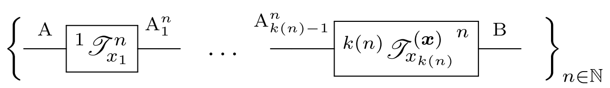
Due to the additional operations enabled by conditioning, it is no longer possible to characterise the generic instruments and transformations of MSOPTs. This is due to the fact that generic conditional instruments do not necessarily take the simple forms given by (22), or, in the deterministic case, (24). However, by introducing an additional assumption, we will demonstrate that the identity transformation remains atomic for every system, even when all conditional instruments are considered.

We recall that the kind of transformations added to an MOPT via closure under conditioning are (procedure 1)



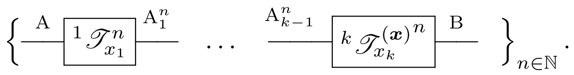
(28)

where $\boldsymbol{x} := x_1, \ldots, x_{k-1}$ and ${}^{i}\mathscr{T}_{x_i}$ are transformations of the original MOPT—thus of the form (22) or limits of Cauchy sequences thereof. We highlight that according to remark 3 the number of conditioning steps we consider is arbitrary but finite (in this case equal to k). Moreover, one has to include transformations that are limits of Cauchy sequences of transformations above,



(29)

Notice that in (29) the systems on the 'inside' wires $\mathrm{A_i}^n$ depend on the index of the sequence, while the input and output systems are fixed. Moreover, also the number of conditioning steps $k(n)$ can vary with the index of the sequence. However, for each n we can take the maximal value $k = \max_m k(m)$ for each sequence item, conditioning with the identity at the extra steps $k - k(n)$ if necessary, thus reducing to consider transformations that are limits of Cauchy sequences of the form



(30)

Clearly (28) can be seen as special case of (30) and analogously the most general instruments to include are Cauchy sequences

$$\left\{\left\{\mathscr{G}^{(\boldsymbol{x})}_{x_k} \circ \mathscr{T}_{\boldsymbol{x}}\right\}^n_{x_k \in \mathsf{X}_k, \boldsymbol{x} \in \mathsf{X}}\right\}_{n\in\mathbb{N}} \subset \mathsf{Instr}\left(\mathrm{A}\right), \tag{31}$$

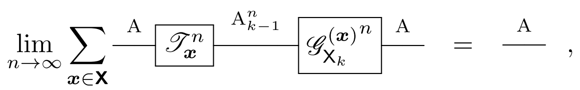
where for simplicity we highlight here only the last conditioning step. Explicitly one has that $\{\{\mathscr{G}^{(\boldsymbol{x})}_{x_k}\}^n_{x_k\in\mathsf{X}_k}\}_{n\in\mathbb{N}} \subset \mathsf{Instr}\left(\mathrm{A}^n_{k-1} \to \mathrm{A}\right)$ and $\{\mathscr{T}^n_{\boldsymbol{x}}\}_{n\in\mathbb{N}} \subset \mathsf{Instr}\left(\mathrm{A} \to \mathrm{A}^n_{k-1}\right)$ with $\boldsymbol{x} := (x_1, \ldots, x_{k-1}) \in \mathsf{X} = \mathsf{X}_1 \times \cdots \times \mathsf{X}_{k-1}$.

7.1. Atomicity of the identity

We can now state and prove the main theorem on the structure of minimal theories with conditioning

Theorem 7. *In every symmetric MSOPT that admits a spanning set of entangled states for every composite system, the identity transformation is atomic for every system.*

Proof. We have to show that the full coarse-graining of every instrument cannot converge to the identity. The only cases we have to check are those of conditional instruments (31), since theorem 6 guarantees the impossibility of decomposing the identity in any other way. The statement then follows proving that the sequence of deterministic transformations obtained by coarse-graining completely over the outcome spaces X and X_k cannot be equal to the identity unless the instruments are trivial decompositions of $\mathscr{I}_\mathrm{A}$, i.e. are of the form $\{p_x\mathscr{I}_\mathrm{A}\}_{x\in\mathsf{X}}$. Thus, we characterize under which conditions the following equality can be satisfied



(32)

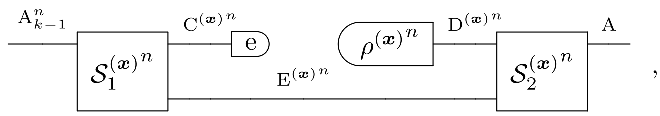
where by $\mathscr{G}^{(\boldsymbol{x})^n}_{\mathsf{X}_k}$ we denote the deterministic transformation given by the full coarse-graining of the test $\{\{\mathscr{G}^{(\boldsymbol{x})}_{x_k}\}^n_{x_k\in\mathsf{X}_k}\}_{n\in\mathbb{N}}$.

The main issue here is that the sequences of transformations in the cascade of instruments above have a complicated dependency of the index n: at an arbitrary conditioning step k' both the input $\mathrm{A}^n_{k'-1}$ and the output $\mathrm{A}^n_{k'}$ are n-dependent thus preventing from straightforwardly applying the stabilization lemma 14. However, we can proceed by iteration from the last step, which has fixed output system A, as follows:

(i) Consider the limit of the coarse-graining of sequence (31). Step by step, starting from the last (kth) conditional gate, each instrument can be coarse-grained to a deterministic map and (32) gives information on their form, especially on the stabilization (dropping of dependence on the sequence step n) of systems involved.

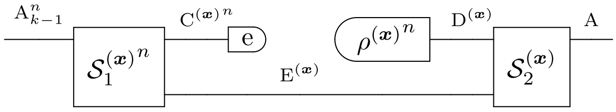
(ii) With the constraints obtained in the last step, we look at the structure of the limit transformations within the conditional instruments. This is crucial to exclude non-trivial decompositions of the identity which holds only if all the limit transformations are proportional to the identity. This second step is quite involved and full details are carried out in appendix D.

Item (i): Thanks to lemma 15 and theorem 5, the deterministic map $\mathscr{G}^{(\boldsymbol{x})^n}_{\mathsf{X}_k}$ is equal to



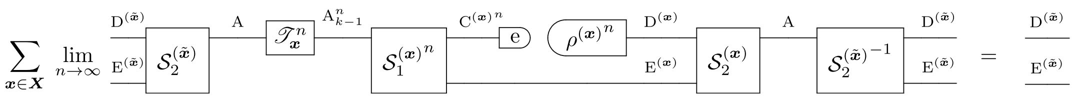
(33)

with $\{\rho^{(\boldsymbol{x})^n}\}_{n\in\mathbb{N}} \subset \mathsf{St}_1\left(\mathrm{D}^{(\boldsymbol{x})^n}\right)$. Contrarily to what happens in lemma 14, the dependence on the index n of the sequence cannot be removed, since the input system A^n_{k-1} also depends on n. However, given that the outcome system A is fixed, it is always possible to find a subsequence where $\mathrm{D}^{(\boldsymbol{x})^n}$, $\mathrm{E}^{(\boldsymbol{x})^n}$, and $\mathcal{S}_2^{(\boldsymbol{x})^n}$ are fixed, being the number of permutations of elementary systems composing A and the ways to group them finite. Hence, one can always consider a subsequence such that $\mathscr{G}_{\mathrm{X}_k}^{(\boldsymbol{x})^n}$ is of the form



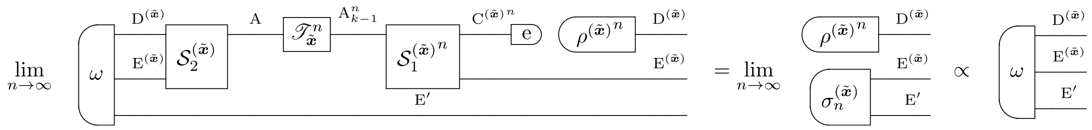
. (34)

The condition expressed in (32) can be rewritten as



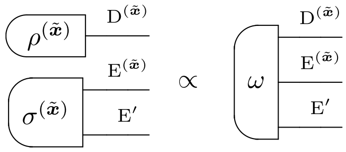
, (35)

where we exploited corollary 1 to exchange the operations of limit and coarse-graining and we have applied the permutation $\mathcal{S}_2^{(\tilde{x})}$ (its inverse) to the left (right) at both sides of the equality. Let now $\omega \in \mathsf{St}\left(\mathrm{D}^{(\tilde{x})}\mathrm{F}^{(\tilde{x})}\right)$ be an atomic entangled state, where $\mathrm{F}^{(\tilde{x})} = \mathrm{E}^{(\tilde{x})}\mathrm{E}'$ for some arbitrary extra ancillary system E'. If we apply both sides of (35) to the state ω, by atomicity it follows that any term of the coarse-graining on the right hand side of (35) must be proportional to ω itself, that is



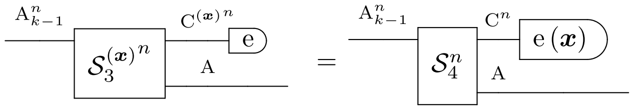
.

Since a Cauchy sequence whose elements are parallel composition of states is such that the sequences of marginals are Cauchy, one can easily prove that the limit is a parallel composition itself (lemma 5), that is



.

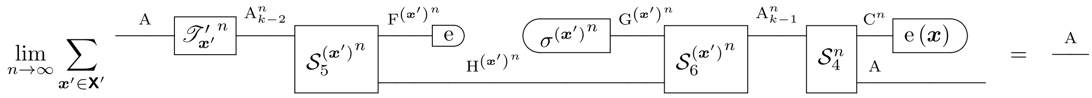
The only possibility for the above equality to hold is that $\mathrm{D}^{(\tilde{x})} = \mathrm{I}$. Since this must hold for every atomic entangled state ω, and the latter states are spanning, we can conclude that $\mathrm{D}^{(\tilde{x})} = \mathrm{I}$ for every $\tilde{\boldsymbol{x}}$, and this implies that (34) is of the form



, (36)

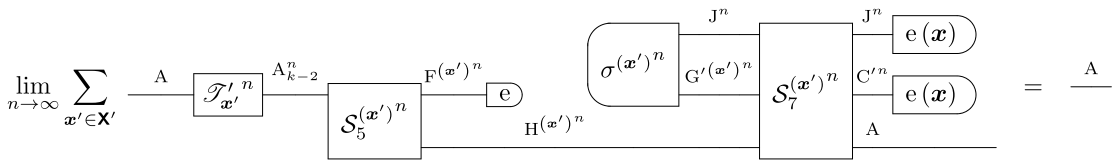
where $\mathcal{S}_3^{(x)^n} := \left(\mathscr{I}_{\mathrm{C}^{(x)^n}} \boxtimes \mathcal{S}_2^{(x)}\right) \circ \mathcal{S}_1^{(x)^n}$. The equality follows reminding that a permutation is completely fixed by how it permutes the elementary subsystems (lemma 12). Therefore, $\mathrm{C}^{(x)^n}$ can be replaced by a fixed system up to a local permutation depending on $\boldsymbol{x}$, which can then be absorbed within the deterministic effect. Clearly the dependence from $\boldsymbol{x}$ in the deterministic effect could be dropped. However, in the study of the transformations composing the above deterministic transformations at Item (ii) the effects in the observation-tests summing to e will depend of the outcome $\boldsymbol{x}$, for this reason we keep it in the notation.

We can now proceeded to make explicit the $k-1$ conditioning step in (32), so we express also the coarse-graining of the instrument at position $k-1$ as a generic deterministic transformation (see lemma 15 and theorem 5) followed by the deterministic transformation at position k derived in (36). The result is of the form

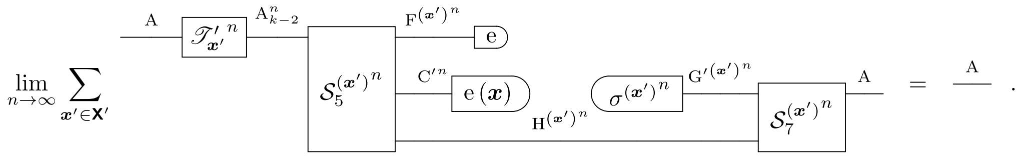


,

where $\boldsymbol{x}' = (x_1, \ldots, x_{k-2}) \in \mathbf{X}' = \mathrm{X}_1 \times \cdots \times \mathrm{X}_{k-2}$, denotes the outcome space of the first $k-2$ instruments. At the first conditioning step (35) we were able to conclude the triviality of system $\mathrm{D}^{(x)}$, here instead we cannot immediately conclude the triviality of system $\mathrm{G}^{(\boldsymbol{x}')}$, but only of part of it. To see this we highlight the fact that $\mathrm{G}^{(\boldsymbol{x}')^n}$ and C^n can share a subsystem J^n made of some elementary subsystems[13]. Let $\mathcal{S}_7^{(\boldsymbol{x}')^n} := \mathcal{S}_4^n \circ \mathcal{S}_6^{(\boldsymbol{x}')^n}$,

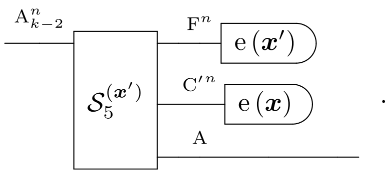


where we used the fact that the deterministic effect is unique and uniquely splits over $\mathrm{J}^n\mathrm{C}'^n$. Exploiting the naturality property of the permutations one can then make the deterministic effects slide to the left obtaining



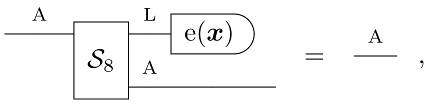
(37)

We can now apply the permutation $\mathcal{S}_7^{(\boldsymbol{x}')}$ (its inverse) on the left (right) side of both circuits, and repeat the argument after (35) to conclude that $\mathrm{G'}^{(\boldsymbol{x}')^n}$ stabilizes to the trivial system. The channel corresponding to last two steps (k and $k-1$) is thus of the form



(38)

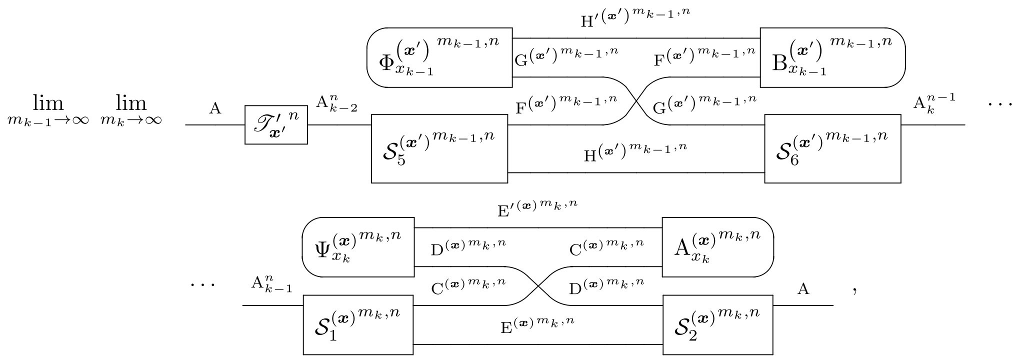
where the dependence of $\boldsymbol{x}'$ disappears in the deterministic effect, but again we keep track of $\boldsymbol{x}'$ since the single events in the observation-test generally depend on it. By iteration, the condition for the coarse-graining in (32) reduces to:



(39)

for a suitable permutation $\mathcal{S}_8$ and system L (notice that by lemma 14 the dependence on n is dropped). The last condition requires the system L to be the trivial system.

Item (ii): Here we check the form of the transformations that appear in the instrument of the sequence (31), and verify that all of them must be proportional to the identity if (32) is satisfied. The idea is to repeat the steps above for transformations instead of their coarse-graining. Since, we are now interested in the single transformations that form the instruments of sequence (31) we consider a fixed outcome $(\boldsymbol{x}', x_{k-1}, x_k) \in \mathbf{X}' \times \mathsf{X}_{k-1} \times \mathsf{X}_k$.

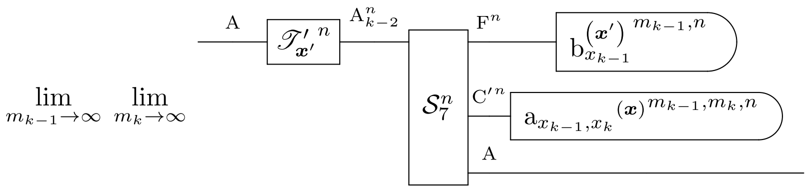


(40)

We highlighted the outcomes of the last two transformations, the kth and the $(k-1)$th ones, in the sequential composition which we write explicitly as limits with respect to the variables m_{k-1} and m_k, respectively. This is due to the fact that the transformations within an instrument of a minimal theory are generally limits of

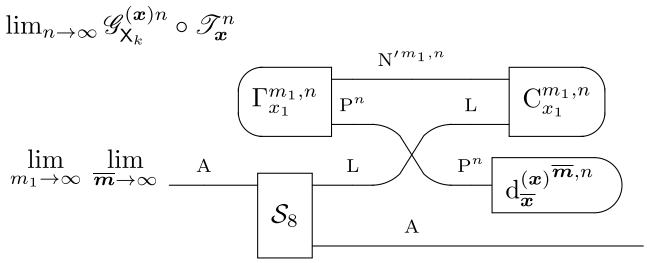
[13] To be precise J^n is defined up to a local permutation. However, this is not important for the discussion since it can be always absorbed within the deterministic effect.

sequences of transformations of the form (22). Following the same argument that led to (38) in the case of transformations (appendix D), one proves that (40) reduces to



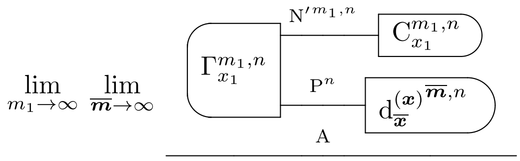
, (41)

where suitable observation-tests appear in place of the deterministic effect of (38). Further iteration leads, in analogy with (39), to the following expression for $\lim_{n\to\infty}\mathscr{G}^{(\boldsymbol{x})n}_{\mathsf{X}_k}\circ\mathscr{T}_{\boldsymbol{x}}{}^n$



, (42)

where $\overline{\boldsymbol{m}}=(m_2,\ldots,m_k)$, $\overline{\boldsymbol{x}}=(x_2,\ldots,x_k)$. Reminding the triviality of system L, we then obtain that any transformation in the instrument (31) must be of the form



, (43)

namely a probability multiplied by the identity map for system A.

Summarizing, we started from a generic Cauchy sequence of instruments of an MSOPT obtained from elementary processes and showed that if it is coarse-grained to the identity then it is of the form

$$\left\{\lim_{m\to\infty} p_x^{m,n}\mathscr{I}_{\mathrm{A}}\right\}_{x\in\mathsf{X}} \qquad \forall n\in\mathbb{N},$$

where $m=m_1,\ldots,m_k$, $x=(x_1,\ldots,x_k)$, and $\mathsf{X}=\mathsf{X}_1\times\cdots\times\mathsf{X}_k$. Since the limit of a Cauchy sequence of probability distributions is still a probability distribution, the limit instrument is equal to

$$\{p'_x\mathscr{I}_{\mathrm{A}}\}_{x\in\mathsf{X}},$$

which concludes the proof. □

7.2. Properties

As in the case of minimal theories, the atomicity of the identity leads to a series of properties of MSOPTs, which persist even after the completion by strong causality.

First, MSOPTs that satisfy the hypothesis of theorem 7 do not admit reversible transformations different from permutations:

Corollary 8. *Every MSOPT satisfying the hypothesis of theorem 7 does not admit of reversible transformations different from permutations.*

The proof follows that of lemma 17, simply noticing that also in this case the argument used in the proof of the theorem 7 guarantees that the only sequence of deterministic transformations that can converge to the identity is the constant one.

Continuing to analyse the properties satisfied by the MSOPTs, from theorem 7 a property that follows, being equivalent to the atomicity of the identity, is NIWD [18]:

Corollary 9. *Every MSOPT satisfying the hypothesis of theorem 7 has NIWD.*

One also has irreversibility, as immediately follows from theorem 3.

Corollary 10. *Every non-trivial MSOPT satisfying the hypothesis of theorem 7 has irreversibility.*

The last feature, which derives directly from lemma 9, is no-broadcasting:

Corollary 11. *Every MSOPT satisfying the hypothesis of theorem 7 does not admit of a broadcasting channel.*

The only property of minimal theories that has not been proved also in the presence of conditional tests is the no-programming theorem (definition 18). This result would require a complete characterisation of the deterministic transformations which is still lacking.

Finally, we would like to remark that if, on the contrary, the hypothesis of theorem 7 are not satisfied, one can have MSOPTs with none of the above properties. The straightforward example is the MSOPT completion of MCT [13], which is the standard CT. This latter result is discussed more in detail at the end of section 9.

8. Classical toy-models

In this section we introduce minimal strongly causal bilocal classical theory (MSBCT), a toy-theory that is locally classical and satisfies strong causality, but still has irreversibility and does not admit a broadcasting channel. The theory is the minimal version of BCT [42] completed under strong causality.

We formalise the theory through the following postulates.

8.1. Postulates of the theory

The first postulate is on the nature of the systems of the theory. We call *classical* any theory such that the set of states of every system has a simplicial basis, all of whose vertices of the simplex belong to a jointly perfectly discriminable set [18]. If all the vertices are jointly perfectly discriminable, this also implies that the theory is convex, every state is proportional to a deterministic one, and, consequently, that the theory is causal [2, 42]. In the latter case the set of states $\mathsf{St}(\mathrm{A})$ is a simplex whose vertices are the states in the perfectly discriminable set and the null state ε_{A}, and $\mathsf{St}_1(\mathrm{A})$ is a simplex itself whose vertices are all those of $\mathsf{St}(\mathrm{A})$ except ε_{A}.

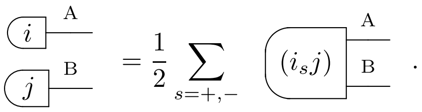
Postulate 1 (Classicality, convexity and system types). The theory Θ is classical and the vertices of state sets are jointly perfectly discriminable. In addition to the trivial system, for every integer $\mathsf{D} \geqslant 1$, $\mathsf{Sys}(\Theta)$ contains a type of system of dimension D.

As observed above, the theory is causal as a consequence of the first postulate. In the following, e_{A} will denote the unique deterministic effect of system A, that amounts to 1 on the vertices of the simplex (except for the vertex given by the null state ε_{A}). The second postulate fixes a parallel composition rule that differs from the tensor product.

Postulate 2 (Parallel composition of systems and states.) For every pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, the dimension of the composite system AB is given by the following rule:

$$\mathsf{D}_{\mathrm{AB}} = \mathsf{D}_{\mathrm{BA}} = \begin{cases} 2\mathsf{D}_{\mathrm{A}}\mathsf{D}_{\mathrm{B}}, & \textit{if } \mathrm{A}, \mathrm{B} \neq \mathrm{I}, \\ \mathsf{D}_{\mathrm{A}}, & \textit{if } \mathrm{B} = \mathrm{I}. \end{cases}$$

Let $\mathrm{I} \neq \mathrm{A}, \mathrm{B}, \mathrm{C} \in \mathsf{Sys}(\Theta)$. Denoting the pure states of the composite system AB as $\mathsf{PurSt}(\mathrm{AB}) = \{(i_-j), (i_+j) \mid 1 \leqslant i \leqslant \mathsf{D}_{\mathrm{A}}, 1 \leqslant j \leqslant \mathsf{D}_{\mathrm{B}}\}$, for all state $i \in \mathsf{PurSt}(\mathrm{A})$, $\mathrm{j} \in \mathsf{PurSt}(\mathrm{B})$ the following parallel composition rule holds:



(44)

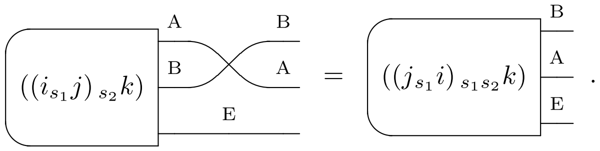
The relation between states of the same composite system defined via compositions in different orders is given by

$$((i_{s_1}j)_{s_2}l) = (i_{s_1}(j_{s_1 s_2}l)), \tag{45}$$

for all indices i, j, k and signs s_1, s_2.

The former postulate also implies that MSBCT is a bilocal theory [42]. In words, this means that it is not possible to discriminate two different states of a composite system by local measurements (and classical communication), as it is instead possible to do in the case of CT or QT which have local discriminability [2, 42, 67, 68]. In the case of bilocal theories, one has generally to perform measurements also on any possible couple of subsystems. For example, considering the state $(i_sj) \in \mathsf{St}(\mathrm{AB})$, with $s = \pm$, by simply measuring the local state of the system $i \in \mathsf{St}(\mathrm{A})$ and $j \in \mathsf{St}(\mathrm{B})$, one would not obtain any information about the sign s. The latter information could be recovered only by performing a measurement on the composite system AB.

Postulate 3 (Swap). The theory is symmetric. Considering $\mathrm{I} \neq \mathrm{A}, \mathrm{B}, \mathrm{E} \in \mathsf{Sys}(\Theta)$, the swap $\mathscr{S}_{\mathrm{A,B}}$ is defined as follows



(46)

Postulate 4 (Preparation- and observation-instruments). Given any system $\mathrm{A} \in \mathsf{Sys}(\Theta)$, a collection $\{\rho_x\}_{x\in\mathsf{X}} \subset \mathsf{St}(\mathrm{A})$ is a preparation instrument if and only if $\sum_{x\in\mathsf{X}} (\mathrm{e}|\,\rho_i)_{\mathrm{A}} = 1$. The observation-instruments of every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ are all the collections $\{\mathrm{a}_y\}_{y\in\mathsf{Y}} \subset \mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$ of generalised effects such that $\{\mathrm{a}_y \boxtimes \mathscr{I}_{\mathrm{E}}\}_{y\in\mathsf{Y}}$ maps preparation instruments of AE to preparation instruments of E for all $\mathrm{E} \in \mathsf{Sys}(\Theta)$.

The sets $\mathsf{Eff}_{\mathbb{R}}(\mathrm{A})$, for every $\mathrm{A} \in \mathsf{Sys}(\Theta)$, are identified by postulate 1 through the property of *joint perfect discriminability* [42].

We highlight that the postulates presented so far are in common with the ones of BCT. The two theories are set apart by the spaces of instruments and transformations.

Postulate 5 (Minimal Strong-causality). The theory Θ is minimal and strongly causal as in definition 19.

Notice that the postulate actually defines a class of theories rather than a single theory. Postulate 5 prescribes only that the theory Θ must satisfy the uniqueness of the decomposition of systems into subsystems (definition 14), without specifying which set of elementary systems is chosen. This aspect itself is not of particular importance, since the results discussed here are valid regardless of this choice. However, for completeness, we will make a couple of remarks in this regard.

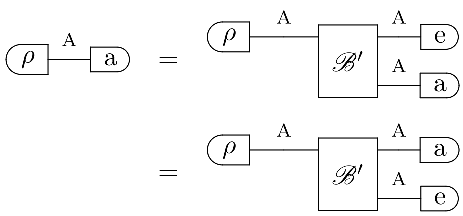
The most conservative choice is that there exists an elementary system of every dimension. This models the possibility of the existence of high-dimensional non-composite systems. For example, in the case of MSBCT, there is no reason to exclude that a system of dimension 8 could be elementary, and not the composition of two systems of dimension 2.

However, it is also possible to make a more minimalist choice in which high-dimensional systems are always seen as compositions of smaller-dimensional ones. In particular, in the case of MSBCT, this would involve defining that the elementary systems of the theory are those of odd dimension, given that starting from them, through the compositional rule described in postulate 2, it becomes possible to reconstruct systems of any dimension. This can be proven by observing that a system of dimension 1 that is not operationally equivalent to the trivial system enables the reconstruction of all even-dimensional systems. Furthermore, the fact that this is the minimal choice can be demonstrated by utilizing the fact that the only way to obtain an odd number as a product of other numbers is for all of those numbers to be odd.

In the following, we write MSBCT as if there were only one theory because the properties that we illustrate actually hold in any of the theories obtained via possible choices of elementary systems.

8.2. Properties

Having introduced the theory, we move on to illustrate its main properties. First of all, it is self-evident that MSBCT is *locally classical*. In other words, if we only consider local operations and measurements on composite systems, the theory cannot be discriminated from CT. For example, local states admit a broadcasting channel. More precisely, forgetting for a moment (9), if the broadcasting channel was defined disregarding the need of copying also arbitrary correlations, i.e. if a map $\mathscr{B}' \in \mathsf{Transf}_1(\mathrm{A}\to\mathrm{AA})$ was defined only requiring that



for any state $\rho \in \mathsf{St}(\mathrm{A})$ and $\mathrm{a} \in \mathsf{Eff}(\mathrm{A})$, then (12) has the required features for a broadcasting channel. However, BCT—that has the same states as MSBCT—features entanglement [11]. There are then data stored in entangled states that would not be broadcast by the above channel $\mathscr{B}'$. In particular, the data that would be destroyed by the above channel are represented by the sign s of pure states $(i_s j)$. Therefore, if one considers the strict definition of broadcasting channel given by (9), then MSBCT does not admit such a map. This descends form lemma 9, since MSBCT admits of systems with dimension greater than 1 whose associated identity transformation is atomic.

In addition to satisfying the no-broadcasting property, MSBCT also has irreversibility of measurement disturbance and it satisfies the property of NIWD. Due to the fact that the entangled states of MSBCT are spanning for the generalised state spaces $\mathsf{St}_{\mathbb{R}}(\mathrm{AB})$ of every composite system AB, MSBCT satisfies the hypothesis of theorem 7 and, consequently, also corollaries 10 and 9.

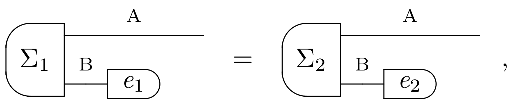
MSBCT provides then an evidence that strong causality and classicality do not prevent no-broadcasting and NIWD. Also, strong causality and classicality do not forbid the existence of 'incompatible' dynamics (irreversibility), despite the fact that all observation-tests of MSBCT are compatible (since their set is equal to that of the observation-tests of BCT which in turn satisfy this property [11]).

We observe that all the properties discussed so far also hold in the case of minimal bilocal classical theory (MBCT), since they only descend from the nature of state spaces along with the fact that the theory is an MOPTs. However, the fact that these properties hold in a minimal classical theory *without* strong causality is not particularly interesting, as they were already been shown in [13], even without bilocal discriminability, through the construction of a minimal version of CT, named MCT.

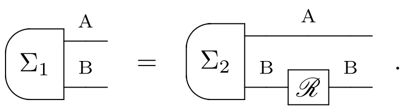
Both in MBCT and in MSBCT no mixed state has a purification, and the operational superposition principle [2, 3] is not satisfied. The latter features are a consequence of the no-go results proven in [42] for general simplicial theories that are not locally tomographic.

Finally, the results concerning the characterization of reversible transformations also allow us to prove that MBCT and MSBCT do not satisfy the property of *essential uniqueness of purification* [2] here reported for the convenience of the reader:

Property 1 (Uniqueness of purification). Let A and B be a couple of generic systems of an OPT. If there exist $\Sigma_1, \Sigma_2 \in \mathsf{PurSt}(\mathrm{AB})$ and $\mathrm{e}_1, \mathrm{e}_2 \in \mathsf{Eff}_1(\mathrm{B})$ such that:



then there exists $\mathscr{R} \in \mathsf{RevTransf}(\mathrm{B} \to \mathrm{B})$ such that:



In a causal theory one would have $\mathrm{e}_1 = \mathrm{e}_2$, corresponding to the unique deterministic effect.

The fact that neither MBCT nor MSBCT satisfy the property just introduced derives from the fact that no permutation in $\mathsf{RevTransf}(\mathrm{B} \to \mathrm{B})$ can map a pure state of the form $(i_s j) \in \mathsf{PurSt}(\mathrm{AB})$ into one of the form $(i'_{s'} j') \in \mathsf{PurSt}(\mathrm{AB})$ whenever, for example, $i \neq i'$. This follows directly from the definition of the swap operation in these theories (postulate 3), since it cannot modify the state of a system on which it is not acting directly.

In conclusion of this section, we observe that the fact that the identity transformation is atomic for every system of MSBCT sets it apart from BCT. In the latter theory, in fact, there always exists a non-trivial test that decomposes the identity [42]. The transformation spaces of MSBCT are *strictly* contained within those of BCT. The same result can also be derived from corollary 8 observing that BCT admits reversible transformations different from permutations.

9. On the relation between theories and their minimal versions

In general, it may not be true that an OPT and its minimal strongly causal version are different. In this respect, we briefly discuss the case of three theories of interest: QT, CT, and BCT.

Let us look first at the quantum case and observe that in this case it is indeed possible to distinguish QT from its minimal strongly causal version. First, being strongly causal [2] QT differs from its minimal version; Minimal Quantum Theory (MQT), which for sure does not admit conditional instruments. By MQT we mean a version of QT where the states and effects spaces are kept unchanged, while the only admissible operations are preparations, observations, the identity and the swap[14]. The procedure to build MQT starting from QT is the same as the one followed to build MCT starting from CT in [13] or MBCT from BCT in section 8. Exploiting corollary 8 one can then show that also the minimal strongly causal version of QT (Minimal Strongly Causal Quantum Theory (MSQT)) differs form QT. Indeed QT admits reversible transformations that are different from permutations, for example the CNOT gate. Therefore, summarizing, in the case of QT the following strict inclusions hold

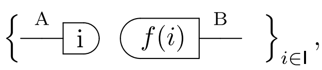
$$\mathsf{Instr}(\mathrm{MQT}) \subset \mathsf{Instr}(\mathrm{MSQT}) \subset \mathsf{Instr}(\mathrm{QT}),$$

where $\mathsf{Instr}(\mathrm{OPT})$ denotes the set of instruments of a particular OPT.

We now move on to the case of BCT. The fact that the identity transformation is atomic for every system of MSBCT makes it different from BCT. In the latter theory, indeed, there always exists a non-trivial test that decomposes the identity [42]. Therefore, the transformations spaces of MSBCT are *strictly* contained within those of BCT. This result can also be derived from corollary 8 observing that BCT admits reversible transformations different from permutations. Moreover, given that MSBCT is a strongly causal version of MBCT, it possess by definition more instruments than MBCT. In summary one then has

$$\mathsf{Instr}(\mathrm{MBCT}) \subset \mathsf{Instr}(\mathrm{MSBCT}) \subset \mathsf{Instr}(\mathrm{BCT}),$$

Last, let us consider the case of standard CT. As anticipated, in the case of MCT, closure by conditioning recovers the original 'full' theory. The reason why a strongly causal version of MCT coincides with CT is that this theory would include all instruments of the form:



for an arbitrary $f: \mathsf{I} \to \mathsf{J}$, where the sets I and J label the pure states of A and B, respectively. Coarse-grainings of the above instruments include all reversible transformation in CT. By the fact that every transformation of CT admits a reversible dilation [2, 11, 42], the result immediately follows. On the other hand, MCT is a strict subset of Minimal Strongly Causal Classical Theory (MSCT) due to the lack of conditional operations in the former. In summary, for CT one has

$$\mathsf{Instr}(\mathrm{MCT}) \subset \mathsf{Instr}(\mathrm{MSCT}) \equiv \mathsf{Instr}(\mathrm{CT}),$$

In this light, the assumption of the existence of entangled states plays a crucial role in the proof of the atomicity of the identity in MSOPTs. Removing this assumption immediately leads to a counterexample to all the statements just made. Indeed, a minimal and strongly causal version of CT coincides with the latter.

As a byproduct we have also proved that the three properties of simpliciality, strong causality and local discriminability are a set of independent properties.

The independence of the three properties derives from the fact that any pair of them does not imply the third one. QT is an example of a theory that is strongly causal and locally tomographic, but non-simplicial. MCT is a locally tomographic non-strongly causal simplicial theory. Finally, MSBCT is a non-locally-tomographic strongly casual simplicial OPT.

Moreover, theories can be exhibited that have exactly one out of the three properties. Indeed, Fermionic quantum theory [9, 43] has strong causality without simpliciality and local discriminability; MBCT has simpliciality without local discriminability and strong causality; finally, MQT has local discriminability without simpliciality and strong causality.

10. Discussion

Exploring the consequences of limiting the allowed dynamics in a generic theory of information processing, we have shown that even a strongly causal CT can consistently support features usually regarded as distinguishing marks of non-classicality. Those are the existence of intrinsically irreversible instruments—namely operations that are not compatible with each other—, NIWD—that is the impossibility of performing non-trivial measurements on a system without perturbing it irreversibly—, and

[14] As for MSBCT, the most conservative approach for choosing the elementary systems in both the minimal and minimal strongly causal versions of QT and CT is to include an elementary system for every possible dimension. Instead, the minimal choice for the set of elementary systems in the case of QT would consist of those whose dimensions are squares of prime numbers—i.e. systems associated with Hilbert spaces of prime dimension. Meanwhile, for CT, the minimal choice would be to consider only systems of prime dimension. Nonetheless, as with for MSBCT, all of the results discussed here hold regardless of the particular choice of the set of elementary systems.

no-broadcasting—which prevents the possibility of 'copying' the state of a system (including its remote correlations). An explicit example of theory with all the above properties has been developed and named MSBCT, showing that they cannot be considered *per se* as signatures of non-classicality.

In order to construct MSBCT, we introduced the class of MSOPTs. These can be seen as MOPTs [13]—i.e. OPTs where all operations are obtained through sequential and parallel composition of preparations, measurements, and permutations—extended by closure under strong causality, thereby ensuring that all conditional instruments are admissible. The minimality criterion introduces the minimal set of dynamical maps consistent with the possible states and measurements of systems, while closure by strong causality enlarges such a minimal set in order to allow for conditioning of instruments on the basis of outcomes of previous experiments. Both MOPTs and MSOPTs aim to model scenarios where operations on a physical system are severely restricted, reflecting the constraints typical of real-world laboratories. However, MSOPTs address a significant limitation of MOPTs: the inability to perform conditioned operations. Such operations are not only routinely possible in experimental settings but are also considered a necessary feature of any physical theory that seeks to accurately describe and model our reality.

The properties of MSBCT have been shown to hold for a broader class of minimal theories. Specifically, all causal symmetric MOPTs and all symmetric MSOPTs that admit a spanning set of entangled states exhibit irreversibility of measurement disturbance, satisfy NIWD, and do not allow broadcasting. These properties arise from the atomicity of the identity transformation, which significantly limits the set of operations that can be performed in these theories on a physical system without causing disturbance.

More specifically, MOPTs do not allow any transformations other than permutations that leave a physical system undisturbed. In the case of MSOPTs, although they may admit transformations that do not disturb the system locally (e.g. (12) in the case of MSBCT), all non-reversible transformations irreversibly destroy correlations with the environment.

Furthermore, the atomicity of the identity transformation also implies that these theories do not admit reversible transformations other than permutations. This property allows, in some cases, the separation of an OPT from its minimal strongly causal version. For example this separation occurs in QT and BCT, but not for CT. While this result may initially seem unremarkable, it provides valuable insights into the information-processing capabilities of CT and QT.

In the classical scenario, any conceivable computation can be achieved through a sequence of measurements followed by conditioned preparations. However, this is not the case in QT, where there are evolutions that cannot be reproduced simply by measuring and re-preparing a system. The fact that CT can be fully recovered from its minimal version (MCT) by simply adding all conditional instruments also highlights the crucial role of entangled states in proving the atomicity of the identity in MSOPTs.

The difference between a MOPT and the associated MSOPT is, instead, always guaranteed by the presence of conditional instruments. For example, instruments of the form (12) are always allowed in MSOPTs and never in MOPTs.

Getting back to the study of the properties of MSBCT, the results presented in [42] guarantee that in this theory no mixed state has a purification, and that superposition of states are not admitted. Furthermore, due to the form of the reversible transformations of the theory, MSBCT also violates essential uniqueness of purification [2].

We also highlight that, even though the notion of classicality adopted by us is related to the simpliciality of the state space, this theory is also Kochen–Specker [69–72] and generalised noncontextual [73–76].

Interestingly, these last few results also hold even when considering just MBCT, i.e. the minimal version of BCT without conditioned instruments.

MSBCT and MBCT also allow us to answer some open questions raised in the conclusions of our previous work [13]. First, they provide another example of theories exhibiting irreversibility despite the full compatibility of the observation-instruments. This proves that MCT is not the only OPT possessing these properties simultaneously. Second, MSBCT shows that adding conditional instruments to a classical theory is not sufficient to guarantee the possibility of generalised broadcasting, highlighting the necessity of the local tomography assumption. This is also the reason why our result does not contradict the one presented in [5, 24].

The construction of MSBCT, along with the minimal versions of CT and QT, also further characterises the relationships between different physical properties, providing a useful contribution to the axiomatisation program of QT. Specifically, we demonstrate that the three properties of simpliciality, strong causality, and local discriminability are pairwise independent (section 9).

As final remarks, we list some topics that in our opinion deserve further investigation.

First, we note that the class of MOPTs enables the construction of OPTs by focusing solely on the state and measurement spaces. This approach is similar to that used in prepare-and-measure scenarios commonly found in the GPT literature. Therefore, given that MOPTs are minimal working examples of OPTs that can be extracted from GPTs defined in the prepare-and-measure scenario, they represent a possible bridge between these two areas of research.

Second, the fact that MOPTs and MSOPTs satisfy the NIWD property suggests the potential for studying whether they could support information-theoretically secure cryptographic protocols for key distribution, similar to QT [56–64]. The strong restrictions on the set of transformations would allow for greater control over all the variables involved in a communication process, thus allowing the problem to be studied in a simplified setting. This would provide interesting insights into the properties that generic information processing theories are required to satisfy to admit intrinsically secure cryptographic protocols. In particular, it would be interesting to determine whether an OPT like MSBCT could admit one, given that it is classical. We hypothesize that this may be feasible by leveraging the fact that an eavesdropper's intervention could be detected, as any attempt to obtain information about the system would irreversibly alter it due to the NIWD property.

In conclusion, we observe that MSOPTs open the possibility of studying the paradigm of *measurement-based computation* [77–83] in an operational setting, given that this computational paradigm consists in carrying out a series of conditional measurements on a suitable initial entangled state. Such a study would open the possibility to further characterize the computational properties of generic operational theories [6, 65]. For example, one could study under which conditions an operational theory is computationally equivalent to its minimal strongly causal version, as in the case of CT, or which additional assumptions are needed.

We finally observe that our findings could have implications concerning [84], where the authors extensively study the relation between the no-broadcasting theorem and contextuality for probabilistic theories. In this scenario, the classical noncontextual theories with no-brodcasting here developed could be used as test toy-models for the implications proved in [84] and for their possible generalization to a broader class of theories.

Data availability statement

No new data were created or analysed in this study.

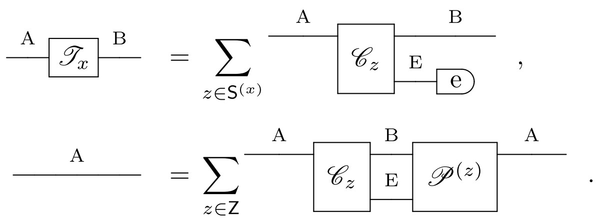
Acknowledgments

DR is currently visiting Perimeter Institute for Theoretical Physics and would like to thank the quantum foundations group there for their hospitality during his visit. This research was supported in part by Perimeter Institute for Theoretical Physics. Research at Perimeter Institute is supported 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. ME acknowledges financial support from the National Science Centre, Poland (Opus project, Categorical Foundations of the Non-Classicality of Nature, Project No. 2021/41/B/ST2/03149). AT acknowledges the financial support of Elvia and Federico Faggin Foundation (Silicon Valley Community Foundation Project ID No. 2020-214365). PP acknowledges financial support from European Union—Next Generation EU through the PNNR MUR Project No. PE0000023-NQSTI.

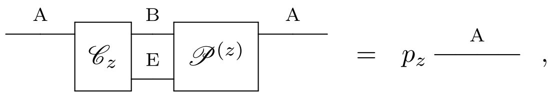
Appendix A. Relation between atomicity of the identity transformation and irreversibility (proof of theorem 3)

Theorem 3. *Let Θ be a causal OPT with a system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ such that $\mathsf{D}_\mathrm{A} \geqslant 2$, and the identity transformation $\mathscr{I}_\mathrm{A}$ to it associated is atomic. Then, there exists an instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$, for some system $\mathrm{B} \in \mathsf{Sys}(\Theta)$ that is intrinsically irreversible. Hence, the theory has irreversibility [13].*

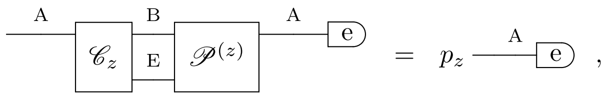
Proof. Suppose by contradiction that this is not the case, i.e. there exists a system A with $\mathsf{D}_\mathrm{A} \neq 1$, where the identity is atomic and there are no intrinsically irreversible instruments having system A as an input. By lemma 10 this implies that whatever instrument $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ one considers it is always possible to find a suitable collection of conditional instruments such that:



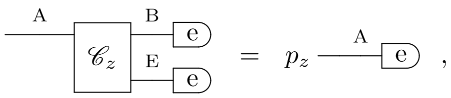
Consequently, since the identity is atomic, it holds that:



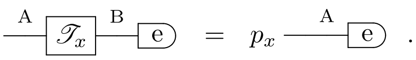
where $\{p_z\}_{z\in\mathsf{Z}}$ is a suitable probability distribution. Applying the deterministic effect on both sides of the above equation obtains:



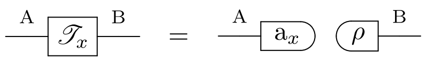
which, given that $\mathsf{P}^{(z)}$ is deterministic $\forall z \in \mathsf{Z}$, implies:



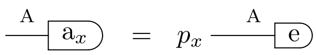
and, consequently, by summing over $z \in \mathsf{S}^{(x)}$ for all $x \in \mathsf{X}$:

 (A1)

A particular set of instruments that satisfies the latter relation consists of those where a fixed deterministic state is prepared following a measurement:

 $\forall x \in \mathsf{X}$,

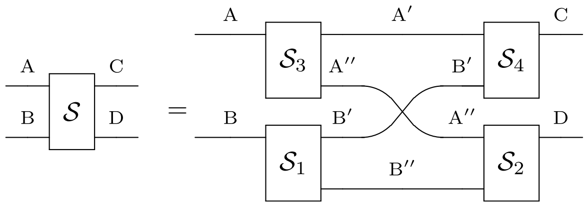
where $\{\mathrm{a}_x\}_{x\in\mathsf{X}} \in \mathsf{Obs}(\mathrm{A})$ is a generic observation-instrument of the theory and $\rho \in \mathsf{St}_1(\mathrm{A})$ is a generic deterministic state. In this particular case, (A1) implies that

 $\forall x \in \mathsf{X}$.

Given that this holds for any observation-instrument of system A, this implies that for this system the only allowed observation-instruments are a randomisation of the deterministic event $\{p_x \mathrm{e}\}_{x\in\mathsf{X}}$. Since a system of this kind has $\dim \mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) = 1$, as can be proven by direct calculation, it must also be $\dim \mathsf{Eff}_{\mathbb{R}}(\mathrm{A}) = \dim \mathsf{St}_{\mathbb{R}}(\mathrm{A}) = \mathsf{D}_\mathrm{A} = 1$, contradicting the hypotheses. □

Appendix B. Characterization of permutations on bipartite systems (proof of lemma 13)

Lemma 13 (General form of permutations on bipartite systems). *In every symmetric OPT with unique decomposition, for any permutation acting on a bipartite system* AB *there exist suitable systems* A′, B′, A′′, B′′, *and reversible transformations* $\mathcal{S}_1$, $\mathcal{S}_2$, $\mathcal{S}_3$, $\mathcal{S}_4$ *such that*



, (20)

where A, B *are generic systems of the theory and* C, D *are systems such that* CD *has the same decomposition in elementary systems as* AB. *In general, any of* A, B, C, D *can be the trivial system, and the same holds also for* A′, A′′, B′, B′′.

Proof. Let us start by considering the ordered decomposition of AB in its elementary subsystems:

$$\{\mathrm{A}_1,\ldots,\mathrm{A}_n,\mathrm{B}_1,\ldots,\mathrm{B}_m\}.$$

The action of $\mathcal{S}$ is to permute its input systems (lemma 12):

$$\{\mathrm{A}_1,\ldots,\mathrm{A}_n,\mathrm{B}_1,\ldots,\mathrm{B}_m\}$$

$$\downarrow\mathcal{S}$$

$$\{\sigma(\mathrm{A}_1),\ldots,\sigma(\mathrm{A}_n),\sigma(\mathrm{B}_1),\ldots,\sigma(\mathrm{B}_m)\}=\{\mathrm{C}_1,\ldots,\mathrm{C}_l,\mathrm{D}_1,\ldots,\mathrm{D}_k\}$$

where, thanks the hypothesis of uniqueness of the decomposition, it holds that $\sigma(\mathrm{A}_1)=C_1$ and so on. If we now define $N=\{1,\ldots,n\},M=\{1,\ldots,m\},L=\{1,\ldots,l\},K=\{1,\ldots,k\}$, the most general transformation that can happen due to the action of $\mathcal{S}$ is that

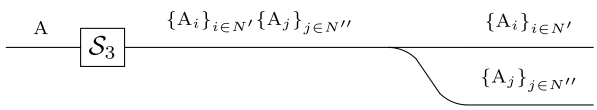
$$\{\mathrm{A}_i\}_{i\in N'}\xrightarrow{\mathcal{S}}\{\mathrm{C}_i\}_{i\in L'},$$
$$\{\mathrm{A}_j\}_{j\in N''}\xrightarrow{\mathcal{S}}\{\mathrm{D}_j\}_{j\in K'},$$

where $N=N'\bigcup N''$, $|N'|=|L'|$, $|N''|=|K'|$, and analogously for B,

$$\{\mathrm{B}_i\}_{i\in M'}\xrightarrow{\mathcal{S}}\{\mathrm{C}_i\}_{i\in L''},$$
$$\{\mathrm{B}_j\}_{j\in M''}\xrightarrow{\mathcal{S}}\{\mathrm{D}_j\}_{j\in K''},$$

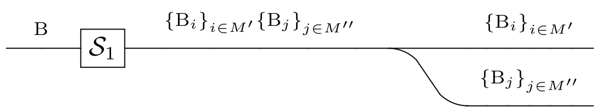
where $M=M'\bigcup M''$, $|M'|=|L''|$, $|M''|=|K''|$, and $L=L'\bigcup L''$, $K=K'\bigcup K''$. With $|S|$ we denote the cardinality of the set S.

Now we want to show that this permutation can always be achieved thorough a transformation with the same form as that of (20). We begin by observing that in the case of system A one can always find a permutation that reorganizes the systems in such a way that the ones that are mapped into states of C are on the top and the ones that are mapped into D are on the bottom,



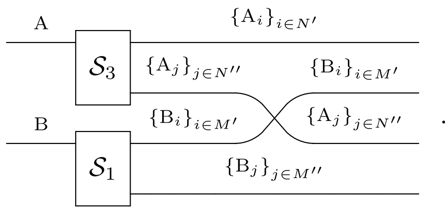
,

where the ordering within N' and N'' is not important. The same holds for B,

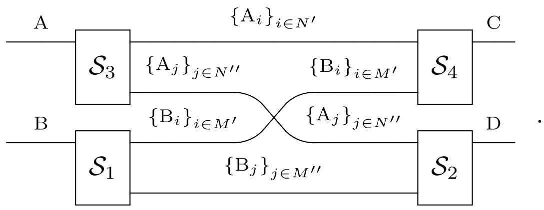


.

Now we have to exchange the subsystems of A that are mapped into D and those of B that are mapped into C. This can be achieved by swapping $\{A_j\}_{j\in N''}$ with $\{B_i\}_{i\in M'}$:



Finally, $\mathcal{S}^2$ and $\mathcal{S}^4$ correctly reorder the subsystems to obtain C and D:

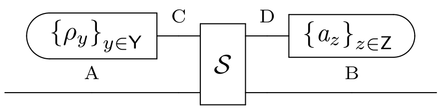


Therefore we have shown that, for any permutation $\mathcal{S}$, it is always possible to find a decomposition as in (20) achieving the same permutation of elementary systems. From the fact that permutations can be completely characterised by how they permute their input systems, the equality between the two transformations follows (lemma 12).

□

Appendix C. Characterization of minimal OPTs' instruments and transformations (proofs of lemma 11 and theorem 4)

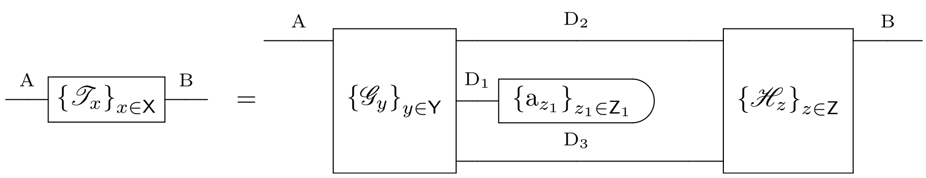
Lemma 11. *In every MOPT any instrument* $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ *obtained as parallel and sequential composition of the elements of (15) is of the form:*



, (17)

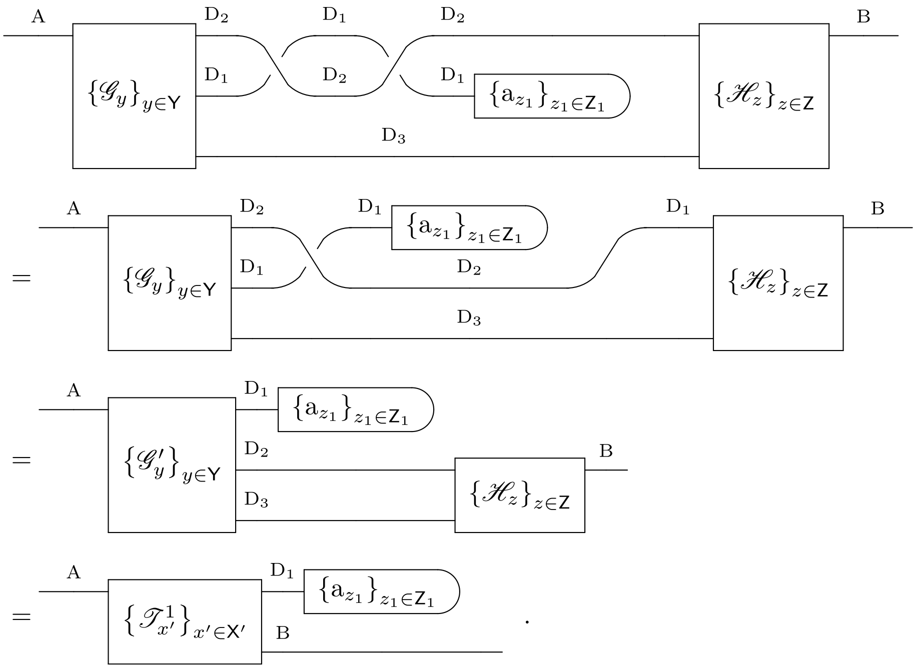
where $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{AC}\to\mathrm{DB})$ *is a suitable braid transformation,* $\{\rho_y\}_{y\in\mathsf{Y}} \in \mathsf{Prep}(\mathrm{C})$, $\{a_z\}_{z\in\mathsf{Z}} \in \mathsf{Obs}(\mathrm{A})$, *the outcome space* $\mathsf{X} = \mathsf{Y}\times\mathsf{Z}$, *and* A, B, C, D $\in \mathsf{Sys}(\Theta)$ *may also be equal to the trivial system [13].*

Proof. The proof of lemma 11 is in a certain way constructive. Let $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ be an instrument that satisfies the hypotheses of the theorem. If this instrument does not contain within its decomposition either a preparation-instrument, or an observation-instrument, then it must be that $\{\mathscr{T}_x\}_{x\in\mathsf{X}} = \mathcal{S}$, i.e. (18) with $\mathrm{C}=\mathrm{D}=\mathrm{I}$. Suppose then that n preparation- and observation-instruments are involved in the decomposition, and let us focus on an observation-instrument—the procedure remains the same if one chooses to start with a preparation-instrument. In the case in which an observation-instrument $\{a_{z_1}\}_{z_1\in\mathsf{Z}_1} \in \mathsf{Obs}(\mathrm{D}_1)$ is present, it is possible to isolate it and rewrite the instrument in the following way:

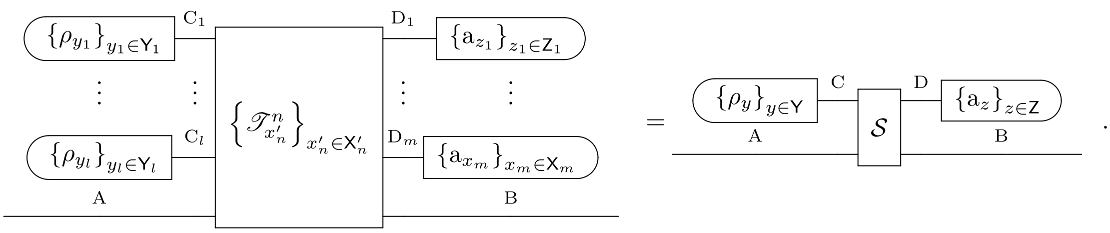


,

where $\{\mathscr{G}_y\}_{y\in\mathsf{Y}}$ and $\{\mathscr{H}_z\}_{z\in\mathsf{Z}}$ are such that $\{\mathscr{T}_x\}_{x\in\mathsf{X}} = \{\mathscr{H}_z\}_{z\in\mathsf{Z}} \circ \left(\mathscr{I}_{\mathrm{D}_2} \boxtimes \{\mathrm{a}_{z_1}\}_{z_1\in\mathsf{Z}_1} \boxtimes \mathscr{I}_{\mathrm{D}_3}\right) \circ \{\mathscr{G}_y\}_{y\in\mathsf{Y}}$ and $\mathrm{D}_1, \mathrm{D}_2, \mathrm{D}_3 \in \mathsf{Sys}(\Theta)$ are suitable systems. Using the reversibility of the braiding, it is possible to write

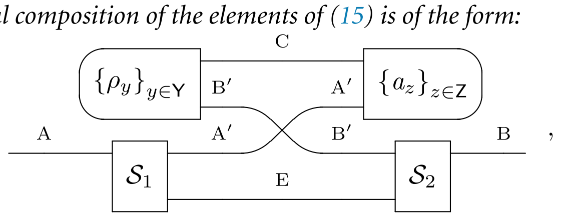


Now it is sufficient to iterate the procedure on $\left\{\mathscr{T}^1_{x'_1}\right\}_{x'_1\in\mathsf{X}'_1}$ until, after n steps, one obtains a transformation $\left\{\mathscr{T}^n_{x'_n}\right\}_{x'_n\in\mathsf{X}'_n} = \mathcal{S}$, which is the searched result.



□

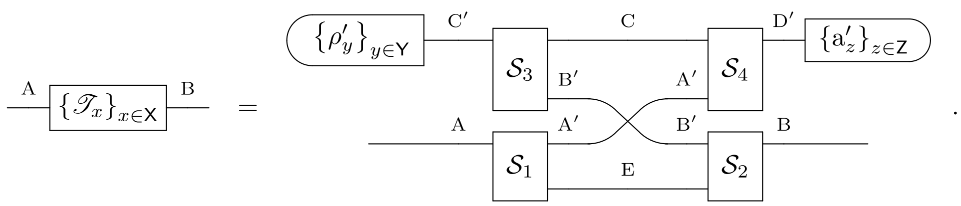
Theorem 4. *In every symmetric MOPT any instrument* $\{\mathscr{T}_x\}_{x\in\mathsf{X}} \in \mathsf{Instr}(\mathrm{A}\to\mathrm{B})$ *obtained as parallel and sequential composition of the elements of (15) is of the form:*



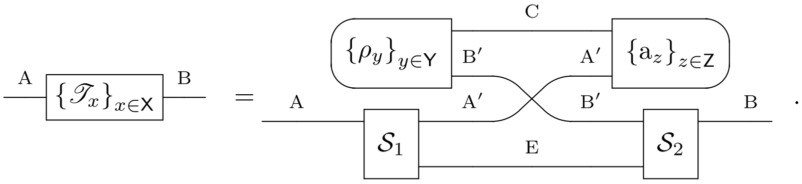
(21)

where $\mathcal{S}_1, \mathcal{S}_2 \in \mathsf{RevTransf}(\Theta)$ *are suitable permutations,* $\{\rho_y\}_{y\in\mathsf{Ys}} \in \mathsf{Prep}(\mathrm{CB}')$, $\{a_z\}_{z\in\mathsf{Z}} \in \mathsf{Obs}(\mathrm{CA}')$, *the outcome space* $\mathsf{X} = Y \times Z$, *and* $\mathrm{A}, \mathrm{B}, \mathrm{A}', \mathrm{B}', \mathrm{C}, \mathrm{E} \in \mathsf{Sys}(\Theta)$ *may also be equal to the trivial system [13].*

Proof. One has just to apply lemma 13 to (17) obtaining



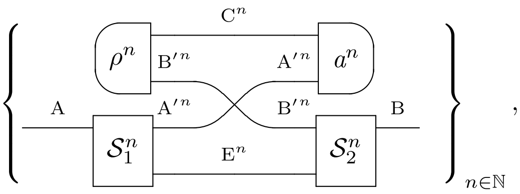
Now absorbing $\mathcal{S}_3, \mathcal{S}_4$ into $\left\{\rho'_y\right\}_{y\in\mathsf{Y}}$ and $\{a'_z\}_{z\in\mathsf{Z}}$ respectively, the proof is concluded:



□

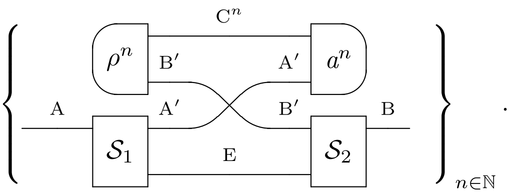
C.1. Characterization of the limits (proofs of lemma 14 and theorem 5)

Lemma 14. *In a symmetric MOPT whenever one considers a Cauchy sequence of generic transformations obtained as parallel and sequential composition of the elements in (16),*



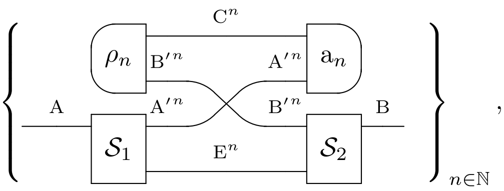
(23)

there always exists a subsequence where the systems E^n, A'^n, B'^n *and the permutations* $\mathcal{S}_1^n$, $\mathcal{S}_2^n$ *are fixed:*



Proof. Let us start by considering a Cauchy sequence such as the one in (23). The proof can be divided in two steps

I. Let us consider the two decompositions in elementary systems of A and B, which we know to be unique (definition 14), and composed at most of a finite number of elementary systems. We remind now that permutations are completely characterised by how they permute the input wires, and since there is only a finite number of ways of permuting a finite number of elements, there must exist at least a pair of permutations $\mathcal{S}^1$ and $\mathcal{S}^2$ that appear infinitely many times together in the decomposition of the elements of the sequence (23). We can now concentrate on the subsequence with this pair of permutations



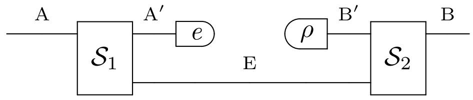
(23)

being (23) Cauchy also its subsequences will be Cauchy and they will have the same limit.

II. We now focus our attention on the systems E_n. Due to the fact that in the previous point we have fixed $\mathcal{S}^1$ and the fact that the decomposition in elementary systems is unique, the systems contained within the composite system $A'_n E_n$ will not change. Therefore, the only thing that can change at the variation of n is how they are grouped. For example if $A'_n = S_1$ and $E_n = (S_2S_3S_4)$, for a different value $n' \neq n$, it must be $A'_{n'} = S_1S_2$ and $E_{n'} = S_3S_4$, or $A'_{n'} = (S_1S_2S_3)$ and $E_{n'} = S_4$, or any other possible regrouping (also the original one) in which the order of the S_i does not change. Given that $A'_n E_n$ can be composed only of a finite number of systems, and analogously for $B'_n E_n$, it is always possible to find at least a system E that appears infinitely many times in the considered sequence. By fixing E, then also the systems A′ and B′ are automatically fixed. This is the searched result.

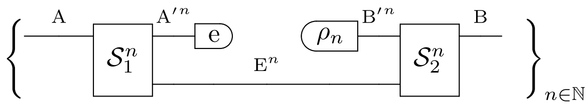
□

Theorem 5. *In a causal symmetric MOPT the limits of Cauchy sequences of deterministic transformations are still of the form [13]*



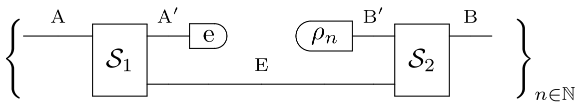
. (24)

Proof. Let us start by considering a Cauchy sequence of deterministic transformations from A to B, which by (24) we know to be of the form



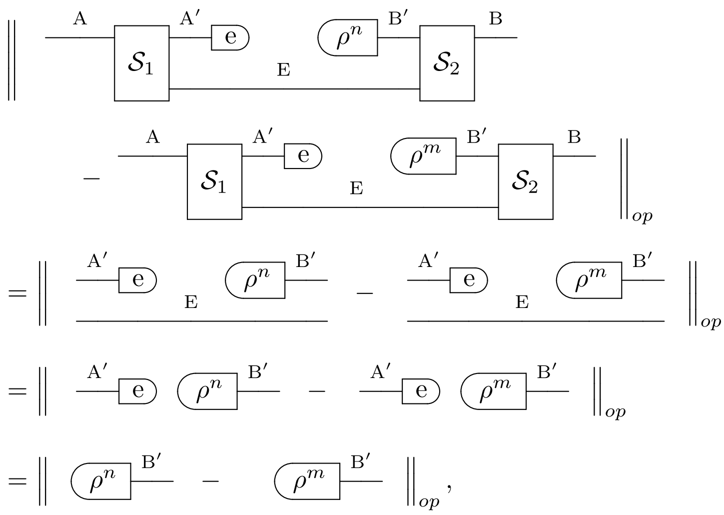
. (C1)

Lemma 14 guarantees the existence of a subsequence where the only elements that can vary thorough the sequence are the states $\{\rho^n\}_{n\in\mathbb{N}} \subset \mathsf{St}_1(\mathrm{B}')$:

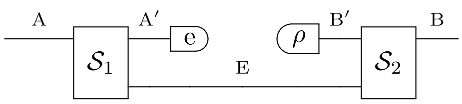


.

The elements of this subsequence satisfy the following inequality relations $\forall n, m \in \mathbb{N}$:



where in the penultimate step lemma 1 was used, while the equality in the last step can be proved using lemma 2. This implies that the sequence of deterministic states of this particular subsequence of (C1) is Cauchy. Therefore, we can conclude that the subsequence considered in this point, and consequently (C1), converges to

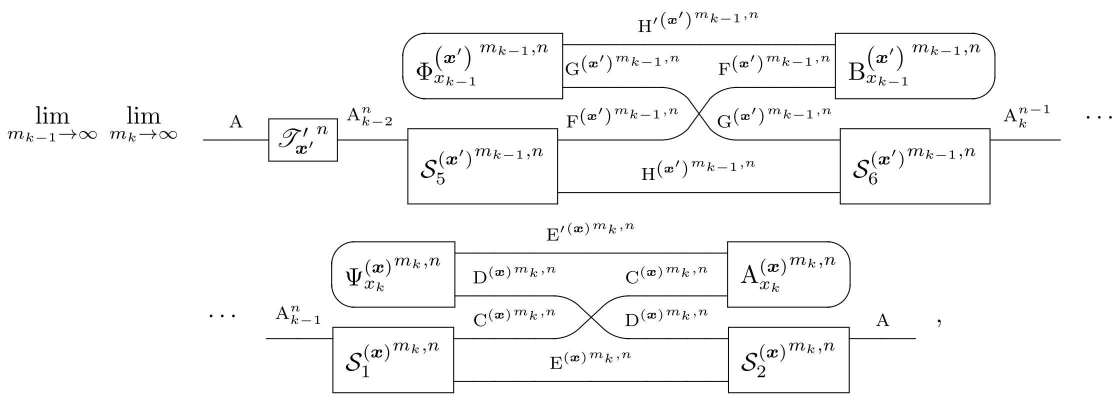


,

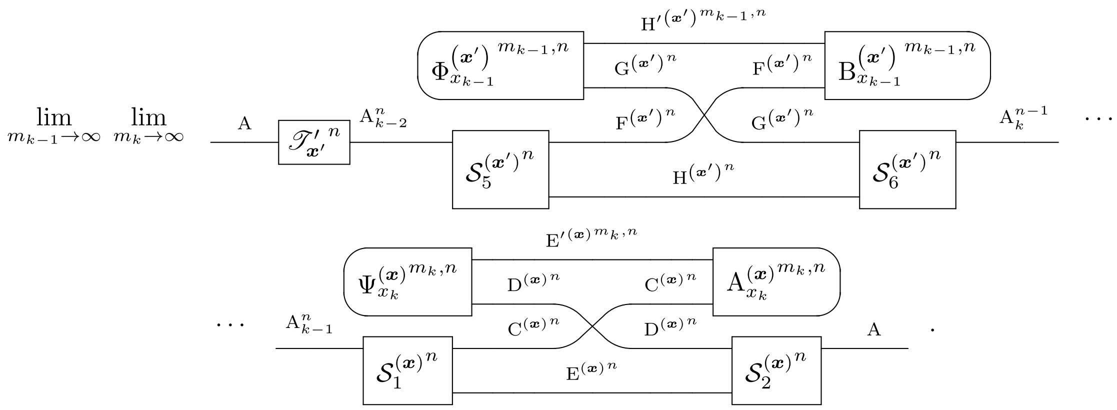
where $\rho = \lim_{n\to\infty} \rho^n$. With this we conclude our proof, since we found the desired result. □

Appendix D. Iteration procedure for transformations in the proof of theorem 7

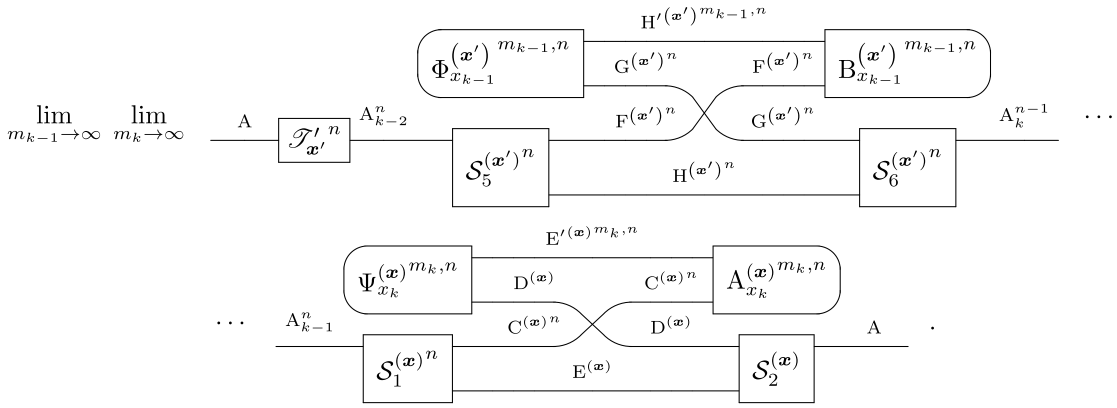
Here we provide the detailed analysis of the structure of single transformations composing the instruments that appear in the sequence (31) and prove that all of them are proportional to the identity if their coarse-graining is the identity itself (32), which means that the identity is an atomic map. Given that we are now interested in studying transformations, in the following discussion we will assume a fixed outcome of the conditional instrument, that is $(\boldsymbol{x}', x_{k-1}, x_k) \in \boldsymbol{X}' \times \mathsf{X}_{k-1} \times \mathsf{X}_k$. We highlighted in the list of outcomes those of the last two transformations in the sequential composition since those are be the ones we will focus on at first. These transformations have generally the form



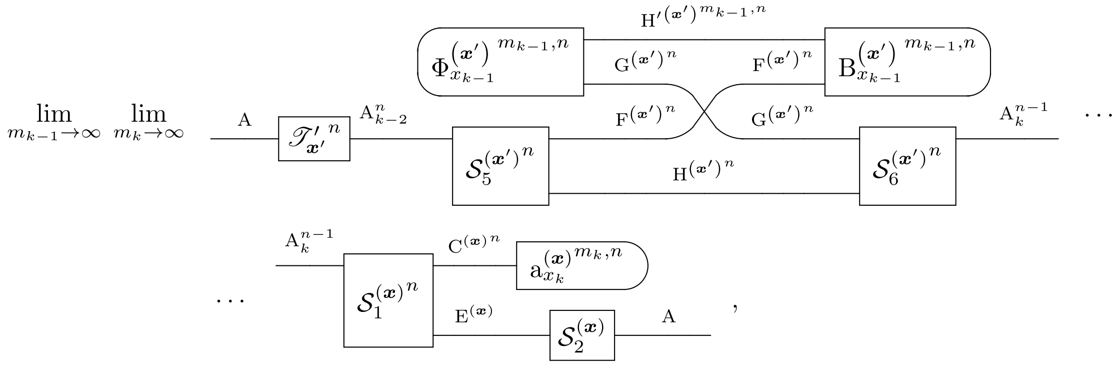
where the limits with respect to the variable m_{k-1} and m_k are due to the fact that the transformations that compose the instruments of (31) are those of MOPTs. Therefore, they may also be limits of sequences of transformations of the form (22). Given lemma 14, one can already remove the dependence of many elements from the indexes m_{k-1} and m_k:



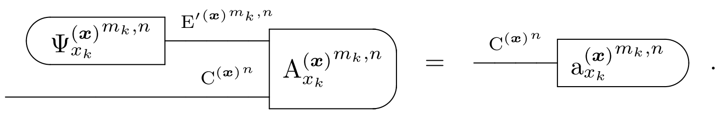
Following the reasoning made for (34), one can the proceed to remove the dependence from the index n on the last transformation in our composition



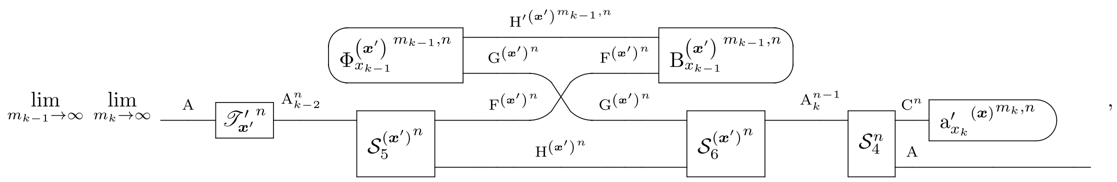
But, we have prove that $\mathrm{D}^{(\boldsymbol{x})} = \mathrm{I}$. Therefore, the transformation becomes



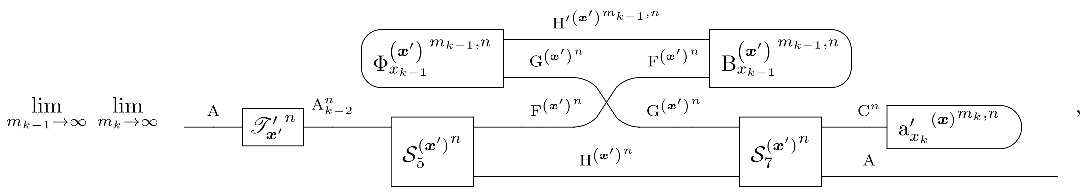
where



The latter equation, following the same passages that lead to (36), becomes

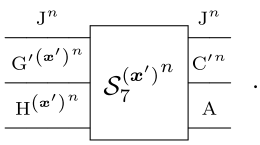


or equivalently



(D1)

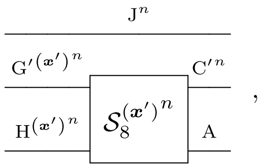
where the local permutation on the system C^n, as done in (36), was absorbed within the effect $\mathrm{a}_{x_k}^{(\boldsymbol{x})^{m_k,n}}$. It is now explicit the reason why before we decided to maintain the dependence form the outcomes for the deterministic effect $\mathrm{e}(\boldsymbol{x})$, acting on C^n. The reason being that while the deterministic effect does not depend on it, the observation-instrument $\left\{\mathrm{a}'_{x_k}{}^{(\boldsymbol{x})^{m_k,n}}\right\}_{x_k\in\mathsf{X}_\mathsf{K}} \in \mathsf{Obs}(\mathrm{C}^n)$ does. We have now arrived at the slightly delicate point exposed before that leads to (37). Here, things are a bit more tricky, since we cannot split the observation-instrument as we did previously for the deterministic effect. However, with a little of patience, also this situation can be dealt with. Let us start again by splitting the systems $\mathrm{G}^{(\boldsymbol{x}')^n} = \mathrm{J}^n\mathrm{G}'^{(\boldsymbol{x}')^n}$ and $\mathrm{C}^n = \mathrm{J}^n\mathrm{C}'^n$[15], and study the the permutation



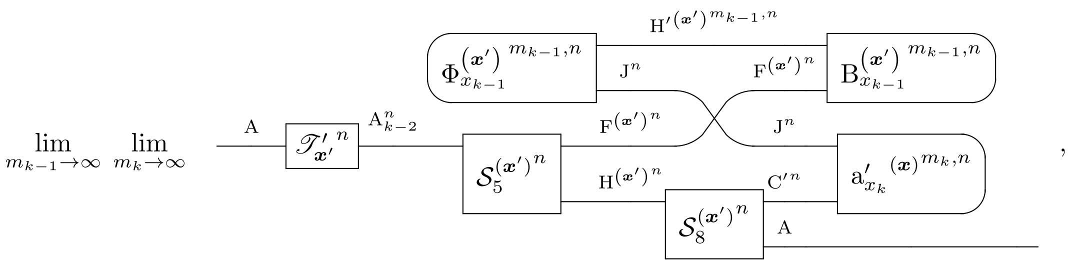
Given that permutations are completely characterized by how they permute their input and output systems (lemma 12), if one finds two different permutations that permute in the same way their input and output

[15] We recall that the system J^n that appears is defined up to a local permutation that can be absorbed within the observation-instrument.

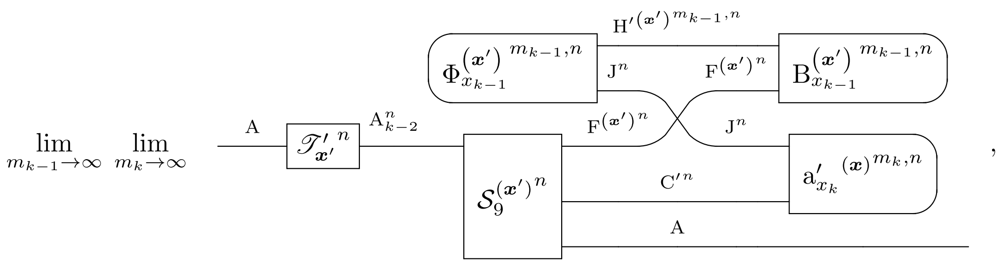
wires, then they are the same transformation. Therefore, an equivalent transformation to $\mathcal{S}_7^{(\boldsymbol{x}')^n}$ is given by



which substituted in (D1) becomes

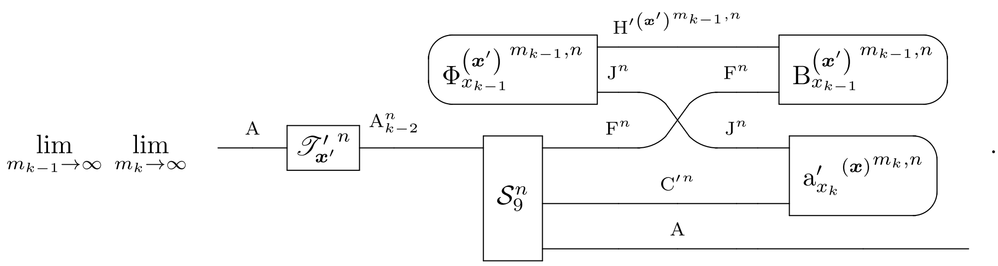


or equivalently,

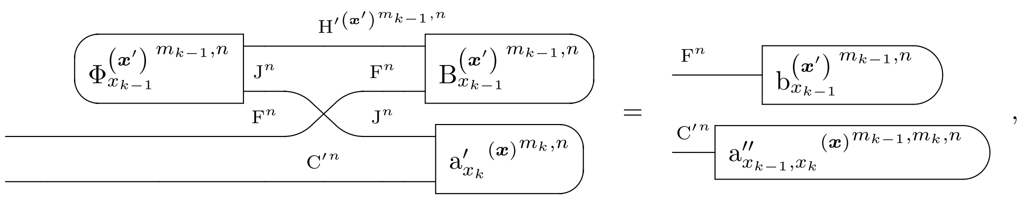


where the fact that it is possible to find a subsequence in which $\mathrm{G}'^{(\boldsymbol{x}')^n} = \mathrm{I}$ was also exploited.

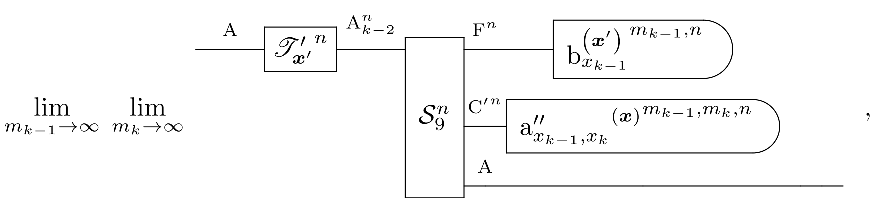
Following the same procedure as before, one can then also remove the dependence from the set of outcomes $\boldsymbol{x}'$ both from the permutation and the system $\mathrm{F}^{(\boldsymbol{x}')^n}$



Exploiting then the fact that

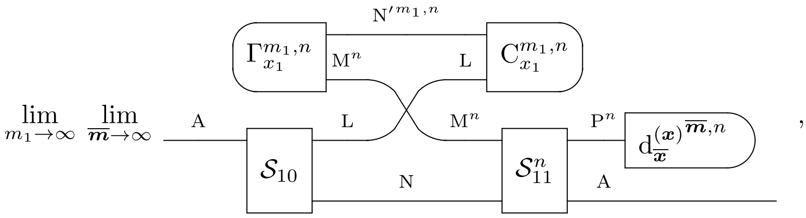


for some suitable observation-instrument, one arrives to

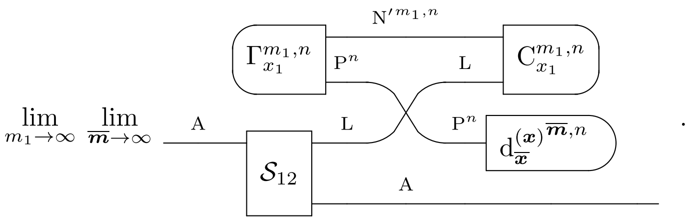


(D2)

which is still of the form (D1), as can be seen by expanding $\mathscr{T}'^{n}_{\boldsymbol{x}'}$ and highlighting the $k-2$th conditional step. So the procedure we followed from (D1) can then be iterated. In particular, by stopping before the last step, what one finds is of the form

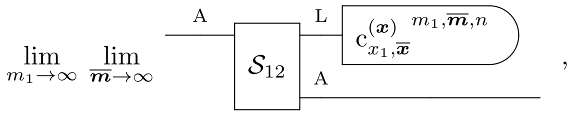


where $\overline{\boldsymbol{m}} = m_2, \ldots, m_k$, $\overline{\boldsymbol{x}} = (x_2, \ldots, x_k)$, $\mathcal{S}^n_{11}$ is a suitable permutation, P^n is a suitable system, $\left\{\mathrm{d}^{(\boldsymbol{x})\overline{\boldsymbol{m}},n}_{\overline{\boldsymbol{x}}}\right\}_{\overline{\boldsymbol{x}}\in\overline{\mathrm{X}}=\mathrm{X}_2\times\cdots\times\mathrm{X}_k}$ is a suitable observation-instrument, the dependence form the index m_1 was already removed from wherever possible and we suppose to consider the subsequence where the leftmost permutation with its systems is fixed. We highlight that the fact that it is possible to remove the dependence form the index n of the sequence from the system L is extremely important to conclude our proof. Since the effect $\mathrm{d}^{(\boldsymbol{x})\boldsymbol{m},n}_{\overline{\boldsymbol{x}}}$ must be completely 'absorbed' by the state $\Gamma^{m_1,n}_{x_1}$ to avoid a transformation which cannot decompose the identity, it must be $\mathrm{M}^n = \mathrm{M}'^n\mathrm{P}^n$ with $\mathrm{M}'^n = \mathrm{I}$, due to the observations made following (35). Therefore, exploiting again the fact that permutations are completely characterized by how they permute the systems wires and grouping whatever is possible, the transformation becomes



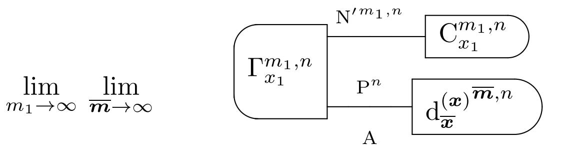
(D3)

We highlight that also $\mathcal{S}_{12}$ still does not depend on n, because it was obtained by composing $\mathcal{S}_{10}$ with a permutation $\mathcal{S} \in \mathsf{RevTransf}(\mathrm{N}\to\mathrm{A})$ and neither N nor A depend form the index n of the sequence. Proceeding now to resolve the calculations related to the state and the effects the following transformation is obtained



which is ill defined unless $\mathrm{L} = \mathrm{I}$.

Finally, substituting this latter result into (D3) one obtains that any transformation that composes the instrument (31) must be of the form



(D4)

if one wants the full coarse-graining of the instrument to be equal to the identity, i.e. for (32) to hold.

ORCID iDs

Davide Rolino https://orcid.org/0009-0007-2192-6178
Marco Erba https://orcid.org/0000-0002-2172-592X
Alessandro Tosini https://orcid.org/0000-0001-8599-4427
Paolo Perinotti https://orcid.org/0000-0003-4825-4264

References

[1] Chiribella G, D'Ariano G M and Perinotti P 2010 Probabilistic theories with purification *Phys. Rev.* A **81** 062348
[2] D'Ariano G M, Chiribella G and Perinotti P 2016 *Quantum Theory From First Principles: An Informational Approach* 1st edn (Cambridge University Press) (https://doi.org/10.1017/9781107338340)
[3] Chiribella G, D'Ariano G M and Perinotti P 2011 Informational derivation of quantum theory *Phys. Rev.* A **84** 012311
[4] Hardy L 1999 Disentangling nonlocality and teleportation (arXiv:quant-ph/9906123)
[5] Barnum H, Barrett J, Leifer M and Wilce A 2006 Cloning and broadcasting in generic probabilistic theories (arXiv:quant-ph/0611295)
[6] Barrett J 2007 Information processing in generalized probabilistic theories *Phys. Rev.* A **75** 032304
[7] Spekkens R W 2007 Evidence for the epistemic view of quantum states: a toy theory *Phys. Rev.* A **75** 032110
[8] Janotta P and Lal R 2013 Generalized probabilistic theories without the no-restriction hypothesis *Phys. Rev.* A **87** 052131
[9] D'Ariano G M, Manessi F, Perinotti P and Tosini A 2014 The Feynman problem and fermionic entanglement: fermionic theory versus qubit theory *Int. J. Mod. Phys.* A **29** 1430025
[10] D'Ariano G M, Manessi F, Perinotti P and Tosini A 2014 Fermionic computation is non-local tomographic and violates monogamy of entanglement *Europhys. Lett.* **107** 20009
[11] D'Ariano G M, Erba M and Perinotti P 2020 Classicality without local discriminability: decoupling entanglement and complementarity *Phys. Rev.* A **102** 052216
[12] Schmid D, Selby J H, Rossi V P, Baldijão R D and Sainz A B 2024 Shadows and subsystems of generalized probabilistic theories: when tomographic incompleteness is not a loophole for contextuality proofs (arXiv:2409.13024)
[13] Erba M, Perinotti P, Rolino D and Tosini A 2024 Measurement incompatibility is strictly stronger than disturbance *Phys. Rev.* A **109** 022239
[14] Perinotti P 2020 Cellular automata in operational probabilistic theories *Quantum* **4** 294
[15] Perinotti P 2021 Causal influence in operational probabilistic theories *Quantum* **5** 515
[16] Busch P 2009 "No information without disturbance": quantum limitations of measurement *Quantum Reality, Relativistic Causality and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (*The Western Ontario Series in Philosophy of Science*) (Springer) pp 229–56 (https://doi.org/10.1007/978-1-4020-9107-0_13)
[17] Heinosaari T, Leppäjärvi L and Plávala M 2019 No-free-information principle in general probabilistic theories *Quantum* **3** 157
[18] D'Ariano G M, Perinotti P and Tosini A 2020 Information and disturbance in operational probabilistic theories *Quantum* **4** 363
[19] Wootters W K and Zurek W H 1982 A single quantum cannot be cloned *Nature* **299** 802–3
[20] Dieks D 1982 Communication by EPR devices *Phys. Lett.* A **92** 271–2
[21] Yuen H P 1986 Amplification of quantum states and noiseless photon amplifiers *Phys. Lett.* A **113** 405–7
[22] Barnum H, Caves C M, Fuchs C A, Jozsa R and Schumacher B 1996 Noncommuting mixed states cannot be broadcast *Phys. Rev. Lett.* **76** 2818–21
[23] Walker T A and Braunstein S L 2007 Classical broadcasting is possible with arbitrarily high fidelity and resolution *Phys. Rev. Lett.* **98** 080501
[24] Barnum H, Barrett J, Leifer M and Wilce A 2007 Generalized no-broadcasting theorem *Phys. Rev. Lett.* **99** 240501
[25] Piani M, Horodecki P and Horodecki R 2008 No-local-broadcasting theorem for multipartite quantum correlations *Phys. Rev. Lett.* **100** 090502
[26] Luo S 2010 On quantum no-broadcasting *Lett. Math. Phys.* **92** 143–53
[27] Chiribella G, D'Ariano G M and Perinotti P 2016 Quantum from principles *Quantum Theory: Informational Foundations and Foils* vol 181, ed G Chiribella and R W Spekkens (Springer) pp 171–221 (https://doi.org/10.1007/978-94-017-7303-4_6)
[28] Hardy L 2001 Quantum theory from five reasonable axioms (arXiv:quant-ph/0101012)
[29] Fuchs C A 2002 Quantum mechanics as quantum information (and only a little more) (arXiv:quant-ph/0205039)
[30] Brassard G 2005 Is information the key? *Nat. Phys.* **1** 2–4
[31] D'Ariano G M 2010 Probabilistic theories: what is special about quantum mechanics? *Philosophy of Quantum Information and Entanglement* 1st edn, ed A Bokulich and G Jaeger (Cambridge University Press) pp 85–126 (https://doi.org/10.1017/CBO9780511676550.007)
[32] D'Ariano G M and Tosini A 2010 Testing axioms for quantum theory on probabilistic toy-theories *Quantum Inf. Process.* **9** 95–141
[33] Masanes L and Müller M P 2011 A derivation of quantum theory from physical requirements *New J. Phys.* **13** 063001
[34] Dakić B and Brukner Č 2011 Quantum theory and beyond: is entanglement special? *Deep Beauty: Understanding the Quantum World Through Mathematical Innovation* ed H Halvorson (Cambridge University Press) pp 365–92 (https://doi.org/10.1017/CBO9780511976971.011)
[35] Barnum H, Barrett J, Leifer M and Wilce A 2008 Teleportation in general probabilistic theories (arXiv:0805.3553)
[36] Barnum H and Wilce A 2011 Information processing in convex operational theories *Electron. Notes Theor. Comput. Sci.* **270** 3–15
[37] Plávala M 2023 General probabilistic theories: an introduction *Phys. Rep.* **1033** 1–64
[38] Coecke B 2006 Kindergarten quantum mechanics: lecture notes *AIP Conf. Proc.* **810** 81–98
[39] Coecke B and Kissinger A 2017 *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge University Press) (https://doi.org/10.1017/9781316219317)
[40] Ludwig G 1985 *Foundations of Quantum Mechanics* (Springer) (https://doi.org/10.1007/978-3-662-28726-2)
[41] Wilce A 2000 Test spaces and orthoalgebras *Current Research in Operational Quantum Logic: Algebras, Categories, Languages* (*Fundamental Theories of Physics*) ed B Coecke, D Moore and A Wilce (Springer) pp 81–114 (https://doi.org/10.1007/978-94-017-1201-9_4)
[42] D'Ariano G M, Erba M and Perinotti P 2020 Classical theories with entanglement *Phys. Rev.* A **101** 042118

[43] Bravyi S B and Kitaev A Y 2002 Fermionic quantum computation *Ann. Phys., NY* **298** 210–26
[44] Lugli M, Perinotti P and Tosini A 2020 Fermionic state discrimination by local operations and classical communication *Phys. Rev. Lett.* **125** 110403
[45] Perinotti P, Tosini A and Vaglini L 2023 Shannon theory for quantum systems and beyond: information compression for fermions *The Quantum-Like Revolution* ed A Plotnitsky and E Bäumer (Springer) pp 135–56 (https://doi.org/10.1007/978-3-031-12986-5_6)
[46] Mac Lane S 1978 *Categories for the Working Mathematician* (*Graduate Texts in Mathematics* vol 5) (Springer) (https://doi.org/10.1007/978-1-4757-4721-8)
[47] Awodey S 2006 *Category Theory* (Oxford University Press) (https://doi.org/10.1093/acprof:oso/9780198568612.001.0001)
[48] Heunen C and Vicary J 2019 *Categories for Quantum Theory: An Introduction* 1st edn (Oxford University Press) (https://doi.org/10.1093/oso/9780198739623.001.0001)
[49] Selby J H, Schmid D, Wolfe E, Sainz A B, Kunjwal R and Spekkens R W 2023 Contextuality without incompatibility *Phys. Rev. Lett.* **130** 230201
[50] Selby J H, Schmid D, Wolfe E, Sainz A B, Kunjwal R and Spekkens R W 2023 Accessible fragments of generalized probabilistic theories, cone equivalence and applications to witnessing nonclassicality *Phys. Rev.* A **107** 062203
[51] Bluhm A, Leppäjärvi L and Nechita I 2024 On the simulation of quantum multimeters (arXiv:2402.18333)
[52] Leppäjärvi L and Sedlák M 2021 Postprocessing of quantum instruments *Phys. Rev.* A **103** 022615
[53] Daffertshofer A, Plastino A R and Plastino A 2002 Classical no-cloning theorem *Phys. Rev. Lett.* **88** 210601
[54] D'Ariano G M, Perinotti P and Tosini A 2022 Incompatibility of observables, channels and instruments in information theories *J. Phys. A: Math. Theor.* **55** 394006
[55] Heinosaari T, Miyadera T and Ziman M 2016 An invitation to quantum incompatibility *J. Phys. A: Math. Theor.* **49** 123001
[56] Bennett C H 1992 Quantum cryptography using any two nonorthogonal states *Phys. Rev. Lett.* **68** 3121–4
[57] Bennett C H, Bessette F, Brassard G, Salvail L and Smolin J 1992 Experimental quantum cryptography *J. Cryptol.* **5** 3–28
[58] Barnett S M and Phoenix S J D 1993 Information-theoretic limits to quantum cryptography *Phys. Rev.* A **48** R5–R8
[59] Ekert A K, Huttner B, Palma G M and Peres A 1994 Eavesdropping on quantum-cryptographical systems *Phys. Rev.* A **50** 1047–56
[60] Kent A 1999 Unconditionally secure bit commitment *Phys. Rev. Lett.* **83** 1447–50
[61] Lo H and Chau H F 1999 Unconditional security of quantum key distribution over arbitrarily long distances *Science* **283** 2050–6
[62] Mayers D 2001 Unconditional security in quantum cryptography *J. ACM* **48** 351–406
[63] Nielsen M A and Chuang I L 2010 *Quantum Computation and Quantum Information: 10th Anniversary Edition* 10th anniversary edn (Cambridge University Press) (https://doi.org/10.1017/CBO9780511976667)
[64] Bennett C H and Brassard G 2014 Quantum cryptography: public key distribution and coin tossing *Theor. Comput. Sci.* **560** 7–11
[65] Barrett J, De Beaudrap N, Hoban M J and Lee C M 2019 The computational landscape of general physical theories *npj Quantum Inf.* **5** 41
[66] Joyal A and Street R 1993 Braided tensor categories *Adv. Math.* **102** 20–78
[67] D'Ariano G M 2006 How to derive the Hilbert-space formulation of quantum mechanics from purely operational axioms *AIP Conf. Proc.* **844** 101–28
[68] Hardy L and Wootters W K 2012 Limited holism and real-vector-space quantum theory *Found. Phys.* **42** 454–73
[69] Bell J S 1966 On the problem of hidden variables in quantum mechanics *Rev. Mod. Phys.* **38** 447–52
[70] Kochen S B and Specker E P 1975 The problem of hidden variables in quantum mechanics *The Logico-Algebraic Approach to Quantum Mechanics: Volume I: Historical Evolution* ed C A Hooker (Springer) pp 293–328 (Originally published in 1967 in J. Math. Mech. 17 59–87) (available at: www.jstor.org/stable/24902153)
[71] Mermin N D 1990 Simple unified form for the major no-hidden-variables theorems *Phys. Rev. Lett.* **65** 3373–6
[72] Mermin N D 1993 Hidden variables and the two theorems of John Bell *Rev. Mod. Phys.* **65** 803–15
[73] Spekkens R W 2005 Contextuality for preparations, transformations and unsharp measurements *Phys. Rev.* A **71** 052108
[74] Schmid D, Selby J H, Pusey M F and Spekkens R W 2024 A structure theorem for generalized-noncontextual ontological models *Quantum* **8** 1283
[75] Schmid D, Selby J H, Wolfe E, Kunjwal R and Spekkens R W 2021 Characterization of noncontextuality in the framework of generalized probabilistic theories *PRX Quantum* **2** 010331
[76] Soltani S, Erba M, Schmid D and Selby J H A local and noncontextual ontological model for bilocal classical theory (forthcoming)
[77] Gottesman D and Chuang I L 1999 Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations *Nature* **402** 390–3
[78] Raussendorf R and Briegel H J 2001 A one-way quantum computer *Phys. Rev. Lett.* **86** 5188–91
[79] Leung D W 2002 Two-qubit projective measurements are universal for quantum computation (arXiv:quant-ph/0111122)
[80] Nielsen M A 2003 Quantum computation by measurement and quantum memory *Phys. Lett.* A **308** 96–100
[81] Raussendorf R, Browne D E and Briegel H J 2003 Measurement-based quantum computation on cluster states *Phys. Rev.* A **68** 022312
[82] Leung D W 2004 Quantum computation by measurements *Int. J. Quantum Inf.* **02** 33–43
[83] Jozsa R 2005 An introduction to measurement based quantum computation (arXiv:quant-ph/0508124)
[84] Jokinen P, Weilenmann M, Plávala M, Pellonpää J, Kiukas J and Uola R 2024 No-broadcasting characterizes operational contextuality (arXiv:2406.07305)