# Measurement incompatibility is strictly stronger than disturbance

Marco Erba ●[*]

*International Centre for Theory of Quantum Technologies, Uniwersytet Gdański, ul. Jana Bażyńskiego 1A, 80-309 Gdańsk, Poland*

Paolo Perinotti ●,[†] Davide Rolino ●,[‡] and Alessandro Tosini ●[§]
*Università degli Studi di Pavia, Dipartimento di Fisica, QUIT Group and INFN Gruppo IV, Sezione di Pavia,*
*via Bassi 6, 27100 Pavia, Italy*

The core of Heisenberg's heuristic argument for the uncertainty principle, involving the famous $\gamma$-ray microscope *Gedankenexperiment*, hinges upon the existence of measurements that irreversibly alter the state of the system on which they are acting, causing an irreducible disturbance on subsequent measurements. The argument was put forward to justify measurement incompatibility in quantum theory, namely, the existence of measurements that cannot be performed jointly—a feature that is now understood to be different from irreversibility of measurement disturbance, though related to it. In this article, on the one hand, we provide a compelling argument showing that measurement incompatibility is indeed a sufficient condition for irreversibility of measurement disturbance, while, on the other hand, we exhibit a toy theory, termed the minimal classical theory (MCT), that is a counterexample for the converse implication. This theory is classical, hence it does not have complementarity nor preparation uncertainty relations, and it is both Kochen-Specker and generalized noncontextual. However, MCT satisfies not only irreversibility of measurement disturbance, but also the properties of no-information without disturbance and no-broadcasting, implying that these cannot be understood *per se* as signatures of nonclassicality.

## I. INTRODUCTION

Since Heisenberg's $\gamma$-ray microscope *Gedankenexperiment* [1], the relation between measurement disturbance and the existence of pairs of observables that cannot be jointly measured has puzzled the authors that tackled quantum measurement theory [2–10]. Over the years, several arguments have been proposed in favor of the fact that these two facets of quantum theory might not necessarily be equivalent. While it may seem intuitively true that the impossibility of jointly measuring two observables necessarily implies measurement disturbance, a proof of this fact has never been given in a theory-independent fashion. On the other hand, it is not even intuitive whether or not the converse implication should hold true. Indeed, no conclusive argument has been given so far in favor or against the latter. The main difficulty in this direction is that quantum theory (QT) exhibits both features, while classical theory (CT) exhibits none of them.

In order to understand the logical relation between incompatibility of observations [11] and irreversibility of measurement disturbance, one needs to move outside the limited scenarios of QT and CT, broadening the perspective to the wider context of general probabilistic theories.

[*]marco.erba@ug.edu.pl
[†]paolo.perinotti@unipv.it
[‡]davide.rolino01@universitadipavia.it
[§]alessandro.tosini@unipv.it

In Ref. [10], the authors exhibit a theory where there are some measurements that cause irreversible disturbance, while the corresponding observations are compatible with all the remaining ones. However, so far no theory has been exhibited such that *all* of its observations are compatible, and yet their measurements cause irreversible disturbance, thus decoupling irreversibility from incompatibility.

In the present article, we address the above question in the framework of operational probabilistic theories (OPTs) [12–15]. These are generic theories of information, including QT and CT as particular cases, but also encompassing a wealth of toy theories sharing the same basic compositional structures for systems and processes. This framework is the appropriate one to seek general arguments about the logical dependency of different properties that physical theories may exhibit. Indeed, the operational-probabilistic framework has been devised in order to survey general physical theories "from the outside." Relevant quantum properties, such as entanglement and contextuality, have been then investigated in a similar generalized scenario. For instance, in Ref. [16] entanglement is established as an inevitable feature of any theory superseding CT while admitting emergent classicality. Furthermore, in Ref. [17] a contextuality witness is deduced in terms of the functional form of an uncertainty relation, thus pinpointing some aspects of quantum uncertainty that may constitute genuine evidence of nonclassicality.

The properties of interest for our analysis are as follows: (i) incompatibility of observations, along with uncertainty, and (ii) irreversibility of measurement disturbance. For case (i), a

family of theories being of interest is the one of *epistemically restricted* classical phase-space theories [18,19], i.e., operational theories with a phase space [20] that is possibly discrete [21], where some restriction in the spirit of Heisenberg's uncertainty principle singles out a minimal volume in phase space that can be identified by a pure state [22]. In the present article, we discuss a family of theories motivated by point (ii), i.e., theories where irreversibility of the measurement disturbance holds.

In detail, in this article we show that the existence of pairs of incompatible observations—a property that we term *incompatibility* for short—is a strictly stronger condition than the existence of operations that irreversibly disturb the state of the system on which they act—that we name *irreversibility*. In order to achieve this, we will prove that the former property implies the latter one, while exhibiting a toy theory that violates the converse implication. The counterexample consists in a theory—that we call minimal classical theory (MCT)—where all observations are compatible, but any measurement irreversibly alters the state of the system. This theory is obtained from CT by restricting to the bone the set of operations one is allowed to perform on a system. The states of systems of MCT being classical, this theory also represents a proof that, contrarily to what is normally believed, the disturbance action caused by the interaction with a system is not a characteristic property of the quantum world. Moreover, this result is complementary to the one of Ref. [17], in that MCT has irreversibility while having no incompatibility of observations—hence no uncertainty thereon—thus being clearly also Kochen-Specker noncontextual. Furthermore, MCT, being embeddable [23] into classical theory, is generalized-noncontextual according to Ref. [24]. We observe that, whereas the pure states of every system can be jointly and perfectly discriminated within the theory, MCT satisfies the property of *no-information without disturbance* [7,10,14,25]. Furthermore, *generalized no-broadcasting* [26,27]—and, as a particular case, *no-cloning*—hold in the theory.

## II. FRAMEWORK

We now sketch the framework of operational probabilistic theories which is here leveraged. An OPT is meant as a theory of systems and their processes. The probabilistic aspect consists in rules to assess the probability of events in any network of processes occurring on a given set of systems. In detail, a generic OPT $\Theta$ has a collection of *systems* and of *tests* thereon. Systems are denoted by capital roman letters A, B, ... $\in \mathsf{Sys}(\Theta)$. As an example, every system in QT corresponds to a complex Hilbert space. Tests, denoted as $\mathsf{T_X} = \{\mathscr{T}_x\}_{x\in\mathsf{X}}$, represent the experiments that one can perform, acting on a given input system A, and obtaining the output system B. The class of tests with input system A and output system B is denoted by $\mathsf{Test}(A \to B)$. Every test consists in a collection of possible *transformations* $\mathscr{T}_x$, labeled by the possible outcomes $x \in \mathsf{X}$ of the experiment. A finite *outcome space* ($\mathsf{X}$) is associated with each test. The class of transformations with input A and output B is denoted by $\mathsf{Transf}(A \to B)$. Referring again to QT, tests are *quantum instruments*, and transformations are *quantum operations*. For example, in a Stern-Gerlach experiment the test that models the action of the

magnetic field is of the form $\mathsf{T_{(\uparrow,\downarrow)}} = \{\mathscr{T}_\uparrow, \mathscr{T}_\downarrow\}$, where the two transformations $\mathscr{T}_\uparrow$ and $\mathscr{T}_\downarrow$ represent the two occurrences in which the system collapses into a state with spin up or down, respectively. A transformation associated with a test whose outcome space has just one element is called *deterministic*. A deterministic transformation does not provide information (the associated test has a unique outcome, occurring with certainty), and can represent, e.g., the evolution of an open system. In QT, a deterministic transformation is a quantum channel.

The first main feature of tests is that they can be performed in a sequence, where a sequence can be defined whenever the input of the subsequent test is the same as the output of the preceding test. Tests (and transformations) will be drawn as boxes, and this makes the representation of a sequence of tests (transformations) more intuitive,

$$\underset{}{\overset{A}{\quad}\boxed{\mathsf{T_X}}\overset{B}{\quad}\boxed{\mathsf{G_Y}}\overset{C}{\quad}} = \underset{}{\overset{A}{\quad}\boxed{\mathsf{G_Y \circ T_X}}\overset{C}{\quad}} \,,$$

where $\mathsf{G_Y \circ T_X} = \{\mathcal{G}_y \circ \mathcal{T}_x\}_{(x,y)\in\mathsf{X}\times\mathsf{Y}}$.

A second defining structure of OPTs is *parallel composition*, that allows one to combine any pair of systems A and B in a *composite system* AB. Given a composite system, moreover, one can independently apply tests $\mathsf{T_X}$ and $\mathsf{G_Y}$ on the two components. The resulting test is the parallel composition $\mathsf{T_X \boxtimes G_Y}$ that is drawn as follows,

$$\begin{array}{c}\overset{A}{\quad}\boxed{\mathsf{T_X}}\overset{C}{\quad}\\[4pt]\overset{B}{\quad}\boxed{\mathsf{G_Y}}\overset{D}{\quad}\end{array} = \overset{AB}{\quad}\boxed{\mathsf{T_X \boxtimes G_Y}}\overset{CD}{\quad} \,,$$

where $\mathsf{T_X \boxtimes G_Y} = \{\mathcal{T}_x \boxtimes \mathcal{G}_y\}_{(x,y)\in\mathsf{X}\times\mathsf{Y}}$. Both sequential and parallel composition are associative.

A special kind of test consists in the *preparation* of a system A. These tests are called *preparation tests*, and their class deserves a dedicated symbol: $\mathsf{Prep}(A)$. The possible transformations of a preparation test are *states*, that can be denoted as $|\rho)_A \in \mathsf{St}(A)$. Similarly, a special class of tests is that representing measurements after which the system A is destroyed, discarded, or just neglected. As mentioned in Ref. [11], these tests are called *observation tests* or *observations* for short, and their set is denoted by $\mathsf{Obs}(A)$. The transformations $(a|_A$ of an observation test are called *effects*, and their set is denoted by $\mathsf{Eff}(A)$. Observation tests are the generalization of positive operator-valued measures (POVMs) of QT to generic theories. We will draw states and effects as

$$\boxed{\rho}\!\!\overset{A}{\quad} \,, \quad \overset{A}{\quad}\!\!\boxed{a} \,,$$

respectively. Preparation (observation) tests can be regarded as special tests whose input (output) system is *trivial*. The (unique) trivial system is denoted by I. From an operational point of view this system represents "nothing the theory cares to describe" [13]. The trivial system behaves as a unit for parallel composition: AI = IA = A.

For every system A of the theory, we require the existence of a deterministic transformation $\mathscr{I}_A$—called *identity*—representing "doing nothing" on the system, i.e., such that $\mathscr{I}_A \mathscr{T} = \mathscr{T} \mathscr{I}_B = \mathscr{T}$, for every transformation $\mathscr{T} \in \mathsf{Transf}(A \to B)$. A transformation $\mathscr{T} \in \mathsf{Transf}(A \to B)$ is

*reversible* if there exists $\mathscr{T}^{-1} \in \mathsf{Transf}(B \to A)$ such that $\mathscr{T}\mathscr{T}^{-1} = \mathscr{I}_B$ and $\mathscr{T}^{-1}\mathscr{T} = \mathscr{I}_A$. Moreover, for every pair of systems A and B, we require the existence of the swap operation representing the exchange of the two systems, i.e.,

$$
\begin{array}{c}
A \qquad\qquad B \\
\diagdown\!\!\diagup \\
B \qquad\qquad A
\end{array} \quad ,
$$

which is a deterministic reversible transformation. Tests and transformations can slide along the crossed wires through a swap.

In an OPT every circuit that starts with a preparation test and ends with an observation test represents a probability distribution for the transformations in the circuit. For example,

$$
\boxed{\rho_x} \!-\!\!^{A}\!\!-\!\boxed{\mathscr{T}_y}\!-\!\!^{B}\!\!-\!\boxed{a_Z} \quad := p\left(x, y, z \mid \rho_X, \mathsf{T}_Y, a_Z\right).
$$

An agent performing a test can discard information regarding the outcome. Correspondingly, in an OPT we require that for every test $\mathsf{T}_X$ and every disjoint partition $\{Z_y\}_{y \in Y}$ of the outcome space $X$ there exists the test $\mathsf{T}'_Y$ representing the same operation, where the outcome $y \in Y$ stands for "the outcome of the test $\mathsf{T}_X$ belongs to $Z_y$." The transformation $\mathscr{T}'_y = \sum_{x \in Z_y} \mathscr{T}_x$ is called *coarse-grained transformation*. Obviously, given a test $\mathsf{T}_X$ the full coarse graining $\mathscr{T}_X = \sum_{x \in X} \mathscr{T}_x$ is deterministic.

The above operational apparatus naturally gives rise to a linear-space structure, where transformations are embedded in real vector spaces.

The spaces of transformations and tests within a theory are required to be Cauchy complete. This follows from the idea that if, within a given theory, there is a procedure to prepare a transformation (or a test) with arbitrary precision, then it is natural to assume that the latter is an ideal transformation (or test) to be included in the theory. This requirement is particularly relevant for the present work, since it allows one to distinguish two theories that share similar building blocks, but where, nonetheless, the operational procedures of one of them cannot be approximated by those of the other one. Importantly, this will allow us to prove that MCT is strictly different from ordinary CT.

For more details about the framework, we refer the reader to Refs. [14,15].

## III. DEFINITIONS

In the following we will consider only *causal* OPTs. These are theories where any system admits of a unique deterministic effect, denoted by $(e|_A \in \mathsf{Eff}(A)$. This condition is equivalent to the property that the probability distributions of preparation tests do not depend on the choice of the observation test at their output—a property known as *no-signaling from the future*. This property also implies spatial no-signaling, namely, the property of *no-signaling without interaction* [14,15].

We now introduce the notion of compatibility of observation tests, which will play a central role in our results. The definition is borrowed from a wide literature on the subject (see, e.g., Refs. [28–30]), where compatibility is ubiquitously identified with joint measurability. In precise terms, we say

that the observation tests $a_X \in \mathsf{Obs}(A)$ and $b_Y \in \mathsf{Obs}(A)$ are *compatible* if there exists a third test $c_{X \times Y} \in \mathsf{Obs}(A)$ such that

$$
\!-\!\!^{A}\!\!-\!\boxed{a_x} = \sum_{y \in Y} \!-\!\!^{A}\!\!-\!\boxed{c_{(x,y)}} \qquad \forall x \in X,
$$

$$
\!-\!\!^{A}\!\!-\!\boxed{b_y} = \sum_{x \in X} \!-\!\!^{A}\!\!-\!\boxed{c_{(x,y)}} \qquad \forall y \in Y.
$$

Accordingly, we will say that a theory has *incompatibility* if it admits of a system A and a pair of observation tests for A that are not compatible.

In order to determine whether an OPT exhibits tests with a disturbance in the sense of Heisenberg, i.e., when an OPT has *irreversibility*, we require the existence of at least a test that irreversibly alters the state of the system on which it acts. In this way, we are stating that these operations set a direction for the arrow of time, in analogy with the second law of thermodynamics. Accordingly, we say that a test is *intrinsically irreversible* if its occurrence precludes the possibility of implementing some other test [31] on the same input system. Notice that in general one can implement a test using ancillary systems, and our definition allows one to postprocess them along with the output system. The precise definition of intrinsic irreversibility is then the following. We say that the test $A_X \in \mathsf{Test}(A \to B)$ is intrinsically irreversible if it *excludes* some other test $B_Y \in \mathsf{Test}(A \to C)$ [31], i.e., there exists a test $B_Y \in \mathsf{Test}(A \to C)$ such that, for every $C_Z \in \mathsf{Test}(A \to BE)$ and every disjoint partition $\{S_x\}_{x \in X}$ of $Z$ with

$$
\!-\!\!^{A}\!\!-\!\boxed{\mathscr{A}_x}\!-\!\!^{B}\!\!-\! = \sum_{z \in S_x} \!-\!\!^{A}\!\!-\!\boxed{\mathscr{C}_z}\begin{array}{c}-\!\!^{B}\!\!-\\ \!-\!^{E}\!\!-\!\boxed{e}\end{array} \tag{1}
$$

there exists no postprocessing $\mathsf{P}_Y^{(z)} \in \mathsf{Test}(BE \to C)$ such that

$$
\!-\!\!^{A}\!\!-\!\boxed{\mathscr{B}_y}\!-\!\!^{C}\!\!-\! = \sum_{z \in Z} \!-\!\!^{A}\!\!-\!\boxed{\mathscr{C}_z}\!-\!^{B}\!\boxed{\mathscr{P}_y^{(z)}}\!-\!\!^{C}\!\!-\! \tag{2}
$$

A first result that we can prove is that a test is intrinsically irreversible if and only if it excludes the identity test. Indeed, if this is the case, the above definition holds choosing $B_Y = \{\mathscr{I}_A\}$. On the other hand, by contradiction, if $A_X$ can be postprocessed to the identity test, then it can be postprocessed to any other test. The detailed proof of the previous statement can be found in Appendix A.

In the light of the above discussion, we will say that a theory has *irreversibility* if it admits of a test that is intrinsically irreversible. Notice that, according to our definition, in QT—where all channels admit of a unitary dilation—no channel is intrinsically irreversible. On the other hand, almost all quantum tests are intrinsically irreversible. Irreversibility thus stems, at least in QT, from the very extraction of information in a measurement.

## IV. INCOMPATIBILITY VERSUS IRREVERSIBILITY

### A. Incompatibility implies irreversibility

We can now prove the first of our two main results: The existence of incompatible observation tests implies that the theory has irreversibility.

Given two observation tests $\{a_x\}_{x\in X}$, $\{b_y\}_{y\in Y} \in \mathsf{Obs}(A)$, if they do not exclude each other, then one can straightforwardly prove that they are compatible. Hence, incompatibility is sufficient for irreversibility. Actually, in any operational theory with nontrivial systems, incompatibility is sufficient also for intrinsic irreversibility of some tests with nontrivial output. In order to prove this, we show that there always exist two tests $\{\mathscr{T}_x\}_{x\in X} \in \mathsf{Test}(A \to B)$ and $\{\mathscr{G}_y\}_{y\in Y} \in \mathsf{Test}(A \to C)$ such that

$$\begin{aligned} \underset{A}{\boxed{a_x}} &= \underset{A}{\boxed{\mathscr{T}_x}}\underset{B}{\boxed{e}} \qquad \forall x \in X, \\ \underset{A}{\boxed{b_y}} &= \underset{A}{\boxed{\mathscr{G}_y}}\underset{C}{\boxed{e}} \qquad \forall y \in Y. \end{aligned} \tag{3}$$

Indeed, in every nontrivial OPT and for every system A, it is always possible to choose a measure-and-prepare test

$$\underset{A}{\boxed{\mathscr{T}_x}}\,^{B} \;=\; \underset{A}{\boxed{a_x}}\;\underset{}{(\rho}\,^{B} \qquad \forall x \in X,$$

where $|\rho)_B \in \mathsf{St}(B)$ is an arbitrary deterministic state of a nontrivial system A, and analogously for $\{b_y\}_{y\in Y}$. We conclude by observing that, if by contradiction either test $\mathsf{T_X}$ or $\mathsf{G_Y}$ does not exclude the other, the observation tests $(e|_B \circ \mathsf{T_X}$ and $(e|_C \circ \mathsf{G_Y}$ are compatible [31], contradicting the hypothesis.

Then, whenever a theory has incompatibility, there must exist at least a pair of tests with nontrivial output that exclude each other, thus being intrinsically irreversible. In summary, *incompatibility of observations implies irreversibility of measurement disturbance*.

### B. Irreversibility does not imply incompatibility

We now proceed to prove the second main result, by exhibiting a toy theory called minimal classical theory (MCT) that has irreversibility but no incompatibility. This theory is obtained by restricting the sets of allowed transformations and tests of CT, while keeping its sets of states and effects untouched. More in detail, the only allowed tests (and consequently transformations) are the ones that can be obtained combining preparation tests and observation tests with the identity and swap operations (and limits of sequences of tests thereof).

MCT is an instance of a broader family of OPTs that can be analogously obtained: Starting from an OPT, one can build its minimal version by only allowing preparation tests and observation tests, permutations of systems, and arbitrary compositions or limits thereof. These theories are called *minimal* OPTs. In Appendix C, the formal definition of this family of theories is presented together with a series of results characterizing their transformations and their properties. As for MCT, the definition and formalization of the results discussed below are presented in Appendix D.

Causal classical theories are here defined as OPTs where the state spaces are simplexes whose vertices (pure states) are jointly perfectly discriminable [15,32]. We can now review some aspects of MCT, actually referring to results that hold for arbitrary minimal OPTs. The tests of a minimal OPT—with

the exclusion of (some of) the limit tests—are of the form



$$\tag{4}$$

where $\{\rho_x\}_{x\in X} \in \mathsf{Prep}(CB')$ and $\{a_y\}_{y\in Y} \in \mathsf{Obs}(CA')$ are generic preparation tests and observation tests, and $\mathcal{S}^{(1)}$ and $\mathcal{S}^{(2)}$ are generic permutations (see Appendix B) [33]. Notice that there is some degree of arbitrariness in the choice of the systems A′, B′, C, E, and in some cases they can be taken as the trivial system I. As a consequence of the realization scheme of tests in Eq. (4), MCT is such that the identity transformation is atomic—i.e., every test whose full coarse graining is equal to $\mathscr{I}$ must be of the form $\{p_x\mathscr{I}\}_{x\in X}$, with $\{p_x\}_{x\in X}$ a probability distribution—for every one of its systems.

The proof of the preceding property proceeds as follows. First, suppose that there is some test that decomposes $\mathscr{I}_A$. This test is the limit of some sequence $\mathsf{T}_X^{(n)}$ of tests of the form (4). The important fact here is that the arbitrary systems $A'_n$, $B'_n$, $E_n$, as well as the permutations $\mathcal{S}_n^{(1)}$ and $\mathcal{S}_n^{(2)}$, for the tests $\mathsf{T}_X^{(n)}$ in the sequence can be taken to be independent of $n$. Then, the full coarse graining of the limit test—that coincides with the limit of the sequence of full coarse grainings—is of the form of Eq. (4) where the observation test $\{a_y\}_{y\in Y}$ reduces to the deterministic effect $(e|_{CA'} \in \mathsf{Eff}(CA')$. Since in our case $A \equiv B$, it must also be $A' \equiv B'$. Moreover, since the overall transformation must be the identity, one can easily check that it must be $\mathcal{S}^{(2)} = [\mathcal{S}^{(1)}]^{-1}$. In summary, one must have



$$\tag{5}$$

and finally, inverting the permutations on both sides, we end up with



which then requires $A' \equiv I$. By the stability of the systems $A'_n$, $B'_n$, $E_n$, also the sequence of tests of the form (4) converging to our decomposition of $\mathscr{I}_A$ must have trivial systems $A' \equiv B' \equiv I$. As a consequence, all such tests must contain transformations proportional to $\mathscr{I}_A$, and so must the limit test.

We finally prove that every theory admitting at least a system of dimension greater than 1 for which the identity transformation is atomic, and MCT as a particular case, has irreversibility. This is shown by contradiction. Suppose that a theory has no intrinsically irreversible tests. Then any test $\mathsf{A_X} \in \mathsf{Test}(A \to B)$ does not exclude the identity and is achievable via a test $\{\mathscr{C}_z\}_{z\in Z} \in \mathsf{Test}(A \to BE)$ such that Eq. (2) holds with $\mathscr{B}_y$ replaced by the identity $\mathscr{I}_A$. Suppose A is a system of dimension greater than one where the iden-

tity transformation is atomic, due to this, one has $(e|_{\mathrm{BE}}\mathscr{C}_z = p_z(e|_{\mathrm{A}}$, which means that the observation test associated to $\mathsf{A}_\mathsf{X} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$ is of the form $\{p_z e\}_{z \in \mathsf{Z}}$, namely it is trivial. Since this is true for every test, all observation tests of system A must be trivial, which is possible only if this system has dimension equal to 1, thus reaching a contradiction. On the other hand, MCT admits systems of dimension greater than 1. Therefore, since for every system of MCT the identity transformation is atomic, it has irreversibility. To conclude the argument, it is sufficient to observe that MCT does not have incompatibility, since it has the same observation tests as CT, where all the observations are compatible. A detailed proof of the previous statement is given in Appendix D 1.

We conclude by observing that the fact that the identity transformation is atomic for any of MCT's systems sets this theory apart from CT. In fact, the latter theory satisfies the opposite property; any system of the theory always admits a test that is a nontrivial decomposition of the identity. Hence MCT's test set is *strictly* contained in CT's one.

## V. DISCUSSION

In this article we have proven that, in a general theory of physical systems, the presence of incompatible observations implies the existence of tests that are intrinsically irreversible, but the reverse does not hold. The counterexample is given in terms of a fully fledged OPT, that we named MCT, whose state spaces are simplexes. Incidentally, in Ref. [34] it is proven that, under the no-restriction hypothesis, the compatibility of all observation tests is equivalent to having simplicial state spaces. At any rate, simpliciality is sufficient for a theory to exhibit full compatibility of observations, and yet, remarkably, it does not preclude the presence of irreversible disturbance, as we have shown here. Indeed, the fact that a theory such as MCT exists is not straightforward.

Notice that it is reasonable to expect that MCT is not the unique theory with full compatibility of observations and irreversibility. Moreover, it is not even clear that such a theory must be simplicial, as suggested, e.g., in Ref. [35], where a possible example is sketched of a theory made of quantum systems whose unique allowed observation test corresponds to an informationally complete POVM.

The notion of intrinsic irreversibility has been also introduced in an operational framework, and characterized as the existence of tests that cannot be postprocessed to the identity—not even with access to arbitrary ancillary systems. The consequent notion of irreversibility—i.e., the property of a theory with an intrinsically irreversible test—is very restrictive, and one may conjecture that it lies at the origin of

thermodynamic irreversibility. The analysis of this hypothesis will be the subject of future studies.

The toy theory presented here, exhibiting irreversibility but also full compatibility of observations, can be used to compare other features which are beyond the scope of this article as well. For example, MCT establishes that classicality is not sufficient for a theory to have full compatibility of tests, where the latter is defined according to Refs. [31,35]. Moreover, MCT satisfies not only no-information without disturbance as follows from the atomicity of the identity transformation [25], but also generalized no-broadcasting [26,27], as follows from the form of the channels of the theory. In particular, as a special case of the latter property, MCT also satisfies no-cloning. Accordingly, MCT represents the evidence that the properties of no-information without disturbance and of no-broadcasting are *not* signatures of nonclassicality *per se*.

We highlight that the fact that MCT satisfies no-broadcasting is not in contradiction with the results of Refs. [26,27]. Indeed, one of the underlying assumptions of the latter works is the possibility of having classical control on outcomes, or, in other words, the possibility of choosing which test to perform conditionally on which outcome has occurred in a preceding test. However, such an assumption was not made in the present work. This establishes that classicality—understood as the joint perfect distinguishability of the pure states—is not in itself a sufficient condition for broadcasting. Now, it would be interesting to determine under which assumptions classicality entails the possibility of generalized broadcasting, e.g., by determining whether the sole addition of the above-mentioned conditional tests would be sufficient [36].

As a final remark, we observe that MCT does not have complementarity nor preparation uncertainty relations, such as Robertson's ones [2], which shows that those are not implied by irreversibility, just as incompatibility is not. It is yet unknown whether or not the converse of the above statement holds.

## APPENDIX A: INCOMPATIBILITY AND INTRINSIC IRREVERSIBILITY

In the present Appendix we provide the formal proof that whenever a test does not exclude the identity, then it does not exclude any other test.

*Lemma 1.* For any two given tests $\mathsf{A}_\mathsf{X} = \{\mathscr{A}_x\}_{x \in \mathsf{X}} \in \mathsf{Test}(\mathrm{A})$ and $\mathsf{B}_\mathsf{Y} = \{\mathscr{B}_y\}_{y \in \mathsf{Y}} \in \mathsf{Test}(\mathrm{A} \to \mathrm{B})$, if $\mathsf{A}_\mathsf{X}$ does not exclude the identity, then it also does not exclude $\mathsf{B}_\mathsf{Y}$.

*Proof.* Writing the nonexclusion relation of $A_X$ with the identity as

$$
\begin{array}{c}
\underline{\quad A\quad}\boxed{\mathscr{A}_x}\underline{\quad A\quad} = \sum_{z\in S_x} \underline{\quad A\quad}\boxed{\mathscr{C}_z}\ \underline{\ A'\ }\boxed{e} \\[4pt]
\underline{\quad A\quad} = \sum_z \underline{\quad A\quad}\boxed{\mathscr{C}_z}\ \underline{\ A'\ }\boxed{\mathscr{P}_y^{(z)}}\ \underline{\ A''\ }\boxed{e},
\end{array}
$$

the result then follows by taking

$$
\begin{array}{c}
\underline{\ A\ }\boxed{\mathscr{B}_y}\underline{\ B\ } = \sum_z \underline{\quad A\quad}\boxed{\mathscr{C}_z}\ \underline{\ A'\ }\boxed{\mathscr{P}_y^{(z)}}\ \underline{\ A''\ }\boxed{\mathscr{B}_y}\underline{\ B\ }\,\boxed{e} \\[8pt]
= \sum_z \underline{\quad A\quad}\boxed{\mathscr{C}_z}\ \underline{\ A'\ }\boxed{\mathscr{P'}_y^{(z)}}\ \underline{\ B\ }\,\boxed{e},
\end{array}
$$

where we defined the new postprocessing $P'^{(z)}_Y := B_Y \circ P^{(z)}_Y \in \mathsf{Test}(AA' \to BA'')$. ∎

## APPENDIX B: PERMUTATIONS AND THEIR PROPERTIES

In the present Appendix we will discuss the particular set of reversible transformations called permutations.

*Definition 1 (Set of permutations).* The *set of permutations*, whose representatives will be indicated with $\mathcal{S}$, is defined as the equivalence class of transformations which are obtained by parallel and sequential composition of swap and identity transformations.

The above defined transformations on bipartite systems satisfy the following characterization theorem:

*Theorem 1 (General form of permutations on bipartite systems).* In every OPT for any permutation acting on a bipartite system, there exist suitable systems A′, B′, A″, B″, and transformations $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \mathcal{S}_4$ such that

$$
\begin{array}{c}
\underline{\ A\ }\boxed{\begin{array}{c}A\ \ C\\ \mathcal{S}\\ B\ \ D\end{array}} = \begin{array}{c}
\underline{\ A\ }\boxed{\mathcal{S}_3}\ \underline{\ A'\ }\ \boxed{\mathcal{S}_4}\ \underline{\ C\ }\\
\underline{\ B\ }\boxed{\mathcal{S}_1}\ \underline{\ B'\ }\ \boxed{\mathcal{S}_2}\ \underline{\ D\ }
\end{array}
\end{array},
$$

$$\tag{B1}$$

where A, B are generic systems of the theory and C, D are systems such that CD is isomorphic to AB.

*Proof.* Let us start by considering the ordered decomposition of AB in the set of subsystems on which $\mathcal{S}$ acts:

$$\{A_1, \ldots, A_n, B_1, \ldots, B_m\}.$$

The action of $\mathcal{S}$ is to permute the order of these subsystems:

$$\{A_1, \ldots, A_n, B_1, \ldots, B_m\}$$

$$\downarrow \mathcal{S}$$

$$\{\sigma(A_1), \ldots, \sigma(A_n), \sigma(B_1), \ldots, \sigma(B_m)\} = \{C_1, \ldots, C_l, D_1, \ldots, D_k\}.$$

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

$$\{A_i\}_{i\in N'} \xrightarrow{\mathcal{S}} \{C_i\}_{i\in L'},$$

$$\{A_j\}_{j\in N''} \xrightarrow{\mathcal{S}} \{D_j\}_{j\in K''},$$

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

$$\{B_i\}_{i \in M'} \xrightarrow{\mathcal{S}} \{C_i\}_{i \in L''},$$

$$\{B_j\}_{j \in M''} \xrightarrow{\mathcal{S}} \{D_j\}_{j \in K''},$$

where $M = M' \bigcup M''$, $\#M' = \#L''$, $\#M'' = \#K''$, and overall $L = L' \bigcup L''$, $K = K' \bigcup K''$. 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 (B1).

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,
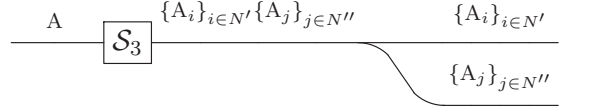


where the ordering of the $A_i$ and $A_j$ is not important. We can then suppose that the same happens also to the subsystems of B,
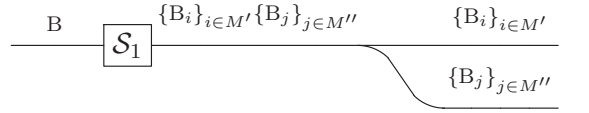


Now we have to take the subsystems of A that are mapped into D and move them down, and vice versa for the ones of B that are mapped into C. This can be achieved by swapping $\{A_j\}_{j \in N''}$ with $\{B_i\}_{i \in M'}$:
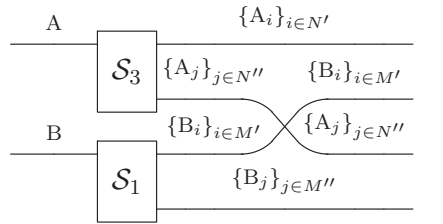


Now to conclude we need only to add two permutations $\mathcal{S}^2$, $\mathcal{S}^4$, that can always be found, to correctly reorder the subsystems to obtain C and D:
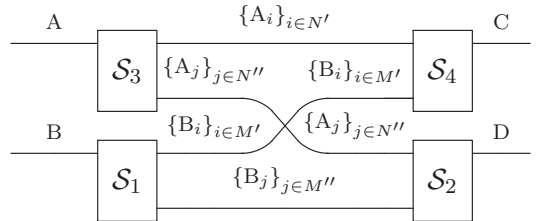


Therefore we have shown that, for any permutation $\mathcal{S}$, it is always possible to find a transformation such as the one in (B1) that permutes the systems as $\mathcal{S}$. From the fact that permutations can be completely characterized by how they permute its input systems, then the equality between the two transformations follows. ∎

*Remark 1.* We highlight that, in general, in the preceding theorem A, B, C, D can be the trivial system and this holds also for A′, A″, B′, B″.

## APPENDIX C: MINIMAL OPERATIONAL PROBABILISTIC THEORIES

*Definition 2 [Minimal operational probabilistic theory (MOPT)].* We define as MOPT an OPT where the only allowed tests are the ones obtainable by composing the elements of



$$(C1)$$

where $\{\rho_i\}_{i \in I}$ and $\{a_j\}_{j \in J}$ are all the possible preparation tests and observation tests of the theory, and the limits of all the Cauchy sequences of tests of this type. Thus the only allowed transformations are those obtainable by sequential and parallel composition

of the elements of

$$
\left\{ \quad \underset{A}{\underline{\quad A \quad}} \quad , \quad \text{(A/B swap)} \quad , \quad \rho\underset{A}{\longmapsto} \quad , \quad \underset{A}{\longmapsto} a \quad \right\} ,
\tag{C2}
$$

for every $A, B \in \mathsf{Sys}(\Theta)$, $|\rho) \in \mathsf{St}(A)$, and $(a| \in \mathsf{Eff}(A)$, and the limits of all the Cauchy sequences of transformations of this type that belong to a test of the theory.

We observe that these are the minimum requirements that can be made on an OPT to cope with the required compositional structure and the Cauchy completeness. In other words, if any of the elements of (C1), or equivalently (C2), or of the limits were removed, the theory could no longer be classified as an OPT.

*Theorem 2.* In every MOPT any transformation $\mathscr{T} \in \mathsf{Transf}(A \to B)$ obtained as a parallel and sequential composition of the elements of (C2) is of the form

$$
\underset{A}{\underline{\quad}}\boxed{\mathscr{T}}\underset{B}{\underline{\quad}} \quad = \quad \text{(diagram with } \rho, a, \mathcal{S}_1, \mathcal{S}_2, \text{ systems } C, B', A', E\text{)} ,
\tag{C3}
$$

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

*Proof.* To prove that this result, we will start by showing that every transformation can be written in the form

$$
\underset{A}{\underline{\quad}}\boxed{\mathscr{T}}\underset{B}{\underline{\quad}} \quad = \quad \rho' \overset{C'}{\longmapsto} \boxed{\mathcal{S}} \overset{D'}{\longmapsto} a' ,
\tag{C4}
$$

Let us consider the decomposition of $\mathscr{T}$ in its constituent elements, (C2), and focus our attention on one of the measurements in it. An effect was chosen, but the procedure remains the same even if one chooses to start with a state. In the case in which neither of them are included in the decomposition it means that $\mathscr{T} = \mathcal{S}$, i.e., (C4) with $C' = D' = I$.

In the case in which an effect $(a_1|$ is present, it is possible to isolate it and rewrite the transformation in the following way,

$$
\underset{A}{\underline{\quad}}\boxed{\mathscr{T}}\underset{B}{\underline{\quad}} \quad = \quad \text{(diagram with } \mathscr{T}_1, a_1, \mathscr{T}_2, \text{ systems } D_1, D_2, D_3\text{)} ,
$$

where $\mathscr{T}_1$ and $\mathscr{T}_2$ are such that $\mathscr{T} = \mathscr{T}_2 \circ (\mathscr{I}_{D_2} \boxtimes (a_1| \boxtimes \mathscr{I}_{D_3}) \circ \mathscr{T}_1$ and $D_1, D_2, D_3 \in \mathsf{Sys}(\Theta)$ are appropriate systems. It is not excluded the possibility of $D_2, D_3$ being the trivial system.

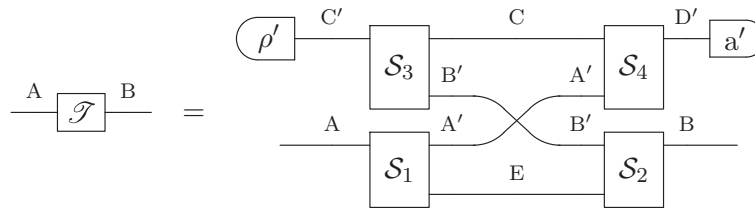Using the reversibility of the permutations, it is possible to write

$$
\text{(sequence of equal diagrams involving } \mathscr{T}_1, \mathscr{T}_2, a_1, \mathscr{T}_1', \mathscr{T}^1 \text{)}
$$

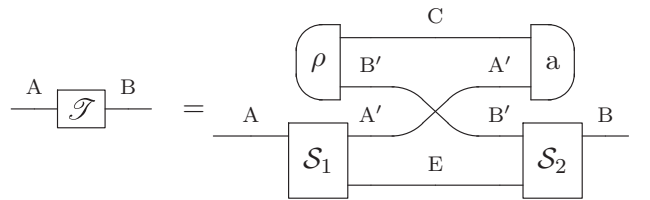Now it is sufficient to iterate the procedure on $\mathscr{T}^1$ until, after $n$ steps, one obtains a transformation $\mathscr{T}^n = \mathcal{S}$. The result would be something of the form
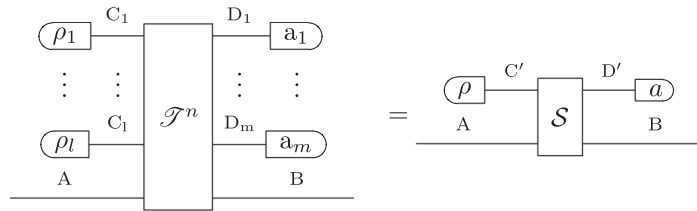


We can now apply Theorem 1 to (C4) obtaining



Now absorbing $\mathcal{S}_3$, $\mathcal{S}_4$ into $|\rho'\rangle_{C'}$ and $(a'|_{D'}$, respectively, the proof is concluded:
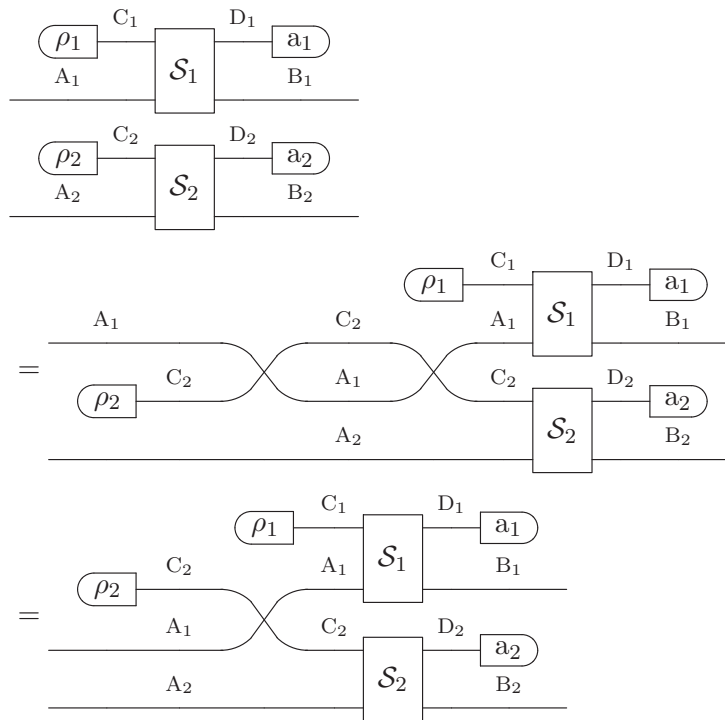


■

One important check for a well-defined OPT is that the spaces of transformations must be closed under parallel and sequential composition. This can be easily proved to hold in every MOPT by exploiting (C4).

Let us start by demonstrating the case of sequential composition:
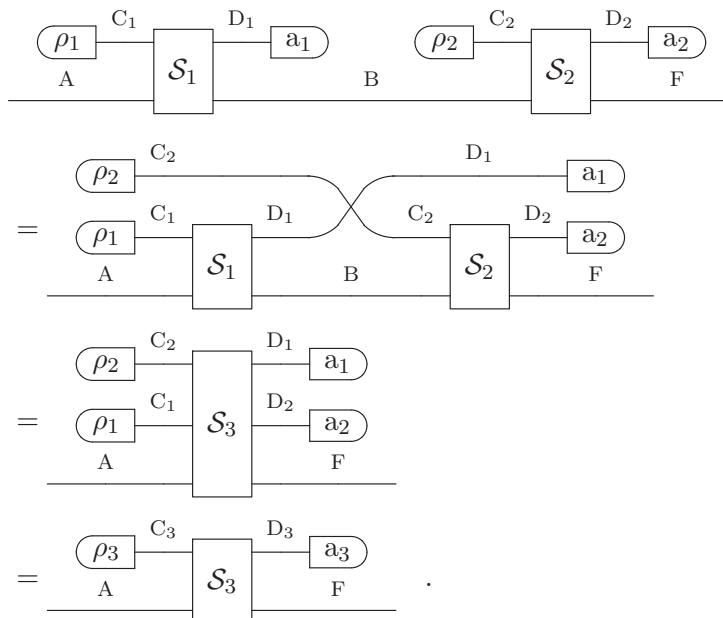
The proof for parallel composition is analogous,



and applying the same procedure on the right-hand side of the circuit with $(a_2|$ one obtains
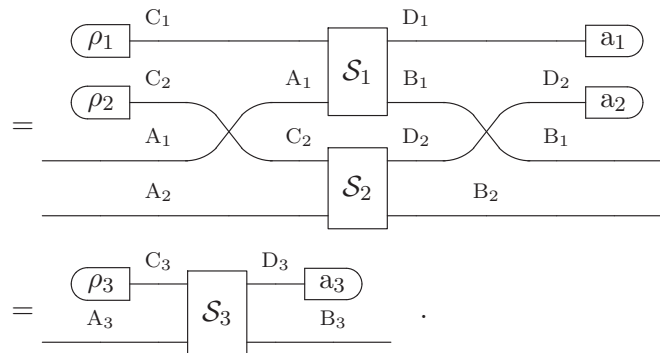


### Properties of MOPTs with the causality assumption

We conclude this Appendix by proving the stability result for the form of the deterministic transformations of MOPTs.

*Lemma 2.* In a causal MOPT every deterministic transformation obtained as composition of the elements in (C2) is of the form



$$\tag{C5}$$
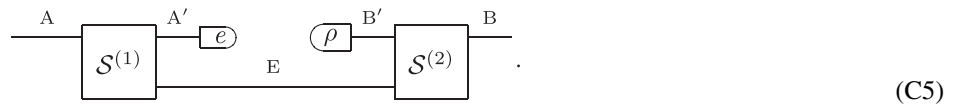
*Remark 2.* The transformation



between the two braid transformations in (C5) is sometimes referred to as "destroy and reprepare," since whatever the input it will "destroy" it and prepare the state $|\rho\rangle$.

*Theorem 3.* In a causal MOPT the limits of Cauchy sequences of deterministic transformations are still of the form (C5).

*Proof.* Let us start by considering a Cauchy sequence of deterministic transformations from A to B, which by (C5) we know to be of the form



$$(C6)$$

The proof can now be subdivided into three steps:

(1) Given that the two systems A and B can only be a composition of a finite number of systems, the sets of permutations that have this systems respectively as input and output [$\mathsf{Permutation}(A \to E)$ and $\mathsf{Permutation}(E' \to B)$ for all appropriate systems $E, E' \in \mathsf{Sys}(\Theta)$] are finite.

Consequently, due to the fact that we have a sequence, i.e., infinite terms, there must exist at least a couple of permutations $\mathcal{S}^1$ and $\mathcal{S}^2$ that appear infinitely many times together "on the outside" of the elements of the sequence (C6). We can now concentrate on the subsequence with this couple of permutations



Since (C6) is a Cauchy sequence, also its subsequences will be Cauchy and they will have the same limit.

(2) We now focus our attention on the systems $E_n$. Due to the fact that in the previous point we have fixed $\mathcal{S}^1$, the systems contained within the composite system $A'_n E_n$ will not change. Therefore, the only change that can occur at the variation of $n$ is how they are grouped.

For example, if $A'_n = S_1$ and $E_n = (S_2 S_3 S_4)$, for a different value $n' \neq n$, it must be $A'_{n'} = S_1 S_2$ and $E_{n'} = S_3 S_4$, or $A'_{n'} = (S_1 S_2 S_3)$ 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. Proceeding exactly as in the previous point we will focus from now on the subsequence where these systems are fixed:



(3) Considering this subsequence we can now easily see that the following relation holds $\forall n, m \in \mathbb{N}$,



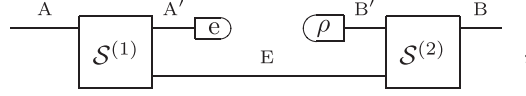where the norm used above is the *operational norm* [14] which has a nice operational interpretation: The distance between two transformations is related to the probability of discriminating them through the best possible procedure one can implement. This norm is well defined over the spaces of transformation since, as observed in the main text, these can be embedded in a real vector space. Furthermore, it satisfies the monotonicity property [14]

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

where $\mathscr{E} \in \mathsf{Transf}(C \to D)$ and $\mathscr{C} \in \mathsf{Transf}(A \to B)$ are deterministic transformations—the equality holds if both $\mathscr{E}$ and $\mathscr{C}$ are reversible—which is what was used in the last steps of the proof.

What this implies is that the sequence of deterministic states of this particular subsequence of (C6) is Cauchy.

We can therefore conclude that the subsequence considered in this point, and consequently (C6), converges to

$$
\begin{array}{c}
\overset{A}{\underline{\phantom{xx}}}\boxed{\mathcal{S}^{(1)}}\overset{A'}{\underline{\phantom{x}}}\boxed{e}\quad\boxed{\rho}\overset{B'}{\underline{\phantom{x}}}\boxed{\mathcal{S}^{(2)}}\overset{B}{\underline{\phantom{xx}}}
\end{array}\;,
$$

where $|\rho)_{B'} = \lim_{n\to\infty}|\rho_n)_{B'}$. With this we conclude our proof, since we found the desired result. ∎

*Theorem 4.* In a causal MOPT whenever one considers a Cauchy sequence of generic transformations obtained as parallel and sequential composition of the elements in (C2),

$$
\left\{
\begin{array}{c}
\rho_n \; B'_n \quad A'_n \; a_n \\
C_n \\
A \quad A'_n \quad B'_n \quad B \\
\mathcal{S}_n^{(1)} \quad E_n \quad \mathcal{S}_n^{(2)}
\end{array}
\right\}_{n\in\mathbb{N}} ,
\tag{C7}
$$

there always exists a subsequence where the systems $E_n$, $A'_n$, $B'_n$ and the permutations $\mathcal{S}_n^{(1)}$, $\mathcal{S}_n^{(2)}$ are fixed:

$$
\left\{
\begin{array}{c}
\rho_n \; B' \quad A' \; a_n \\
C_n \\
A \quad A' \quad B' \quad B \\
\mathcal{S}^{(1)} \quad E \quad \mathcal{S}^{(2)}
\end{array}
\right\}_{n\in\mathbb{N}} .
$$

*Proof.* The proof of this result consists in going over the first two points of the proof of Theorem 3, and applying them to the case considered here. ∎

## APPENDIX D: MINIMAL CLASSICAL THEORY

We will now discuss in detail minimal classical theory (MCT). The procedure for the construction of generic OPTs presented in Ref. [15] guarantees that the postulates here presented are sufficient to to construct a well-defined operational theory.

*Postulate 1 (Classicality, convexity, and type of systems).* The theory $\Theta$ is classical, convex, and satisfies local discriminability. In addition to the trivial system, for every integer $D \geqslant 1$, $\mathsf{Sys}(\Theta)$ contains a type of system of size $D$.

*Postulate 2 (Preparation and observation tests).* Given any system $A \in \mathsf{Sys}(\Theta)$, a collection $\{\rho_x\}_{x\in\mathsf{X}} \subset \mathsf{St}(A)$ is a preparation test if and only if $\sum_{x\in X}(e|\rho_x)_A = 1$. The observation tests of every system $A \in \mathsf{Sys}(\Theta)$ are all the collections $\{a_y\}_{y\in\mathsf{Y}} \subset \mathsf{Eff}_{\mathbb{R}}(A)$ of generalized effects such that $\{(a_y|_A \boxtimes \mathscr{I}_E\}_{y\in\mathsf{Y}} \subset \mathsf{Eff}_{\mathbb{R}}(AE)$ maps preparation tests of AE to preparation tests of E for all $E \in \mathsf{Sys}(\Theta)$.

Where $\mathsf{Eff}_{\mathbb{R}}(A)$ for every $A \in \mathsf{Sys}(\Theta)$ is defined by Postulate 1 through the property of joint perfect discriminability.

*Postulate 3 (Transformations and tests).* The only allowed tests are the ones given by the composition of the elements of

$$
\left\{
\begin{array}{ccccc}
\overset{A}{\underline{\phantom{xx}}} & , & \underset{B}{\overset{A}{\times}}\underset{A}{\overset{B}{}} & , & \boxed{\{\rho_i\}_{i\in\mathsf{I}}}\overset{A}{\underline{\phantom{x}}} & , & \overset{A}{\underline{\phantom{x}}}\boxed{\{a_j\}_{j\in\mathsf{J}}}
\end{array}
\right\}
\tag{D1}
$$

where $\{\rho_i\}_{i\in\mathsf{I}}$ and $\{a_j\}_{j\in\mathsf{J}}$ are all the possible preparation tests and observation tests allowed in the theory by Postulate 2, and the limits of all Cauchy sequences of tests of this type. Thus the only allowed transformations are the ones obtainable by sequential and parallel composition of the elements of

$$
\left\{
\begin{array}{ccccc}
\overset{A}{\underline{\phantom{xx}}} & , & \underset{B}{\overset{A}{\times}}\underset{A}{\overset{B}{}} & , & \boxed{\rho}\overset{A}{\underline{\phantom{x}}} & , & \overset{A}{\underline{\phantom{x}}}\boxed{a}
\end{array}
\right\}
\tag{D2}
$$

for every $A, B \in \mathsf{Sys}(\Theta)$, $|\rho) \in \mathsf{St}(\Theta)$, and $(a| \in \mathsf{Eff}(\Theta)$, and the limits of all the Cauchy sequences of events of this type that belong to a test of the theory.

We recall the following definitions.

*Definition 3 (Convex OPT).* An OPT $\Theta$ is *convex* if $\mathsf{St}(A)$ coincides with its convex hull for all $A \in \mathsf{Sys}(\Theta)$.

*Definition 4 (Local discriminability).* It is possible to discriminate any pair of states of composite systems using only local measurements [14].

On top of this, we remind that this theory is causal since every state is proportional to a deterministic one [14].

### 1. MCT has full compatibility of observation tests

We will now prove that MCT satisfies the property of full compatibility of observation tests. Let us consider two of them. The most generic form they can take in an $n$-dimensional system is the following:

$$\{a_x\}_{x \in \mathsf{X}} = \{p_0^0(0| + p_1^0(1| + \cdots + p_n^0(\mathsf{n}|, p_0^1(0| + p_1^1(1| + \cdots + p_n^1(n|, \ldots, p_0^m(0| + p_1^m(1| + \cdots + p_n^m(\mathsf{n}|\} \in \mathsf{Eff}(A)$$

$$\{b_y\}_{y \in \mathsf{Y}} = \{q_0^0(0| + q_1^0(1| + \cdots + q_n^0(\mathsf{n}|, q_0^1(0| + q_1^1(1| + \cdots + q_n^1(n|, \ldots, q_0^k(0| + q_1^k(1| + \cdots + p_n^k(\mathsf{n}|\} \in \mathsf{Eff}(A),$$

where $\sum_{i=0}^m p_j^i = 1 \quad \forall j = 0, \ldots, n$ with $p_j^i \in [0, 1] \forall i, j$ and analogously for $q_j^i$. Defining now

$$\{c_{(i,j)}\}_{(i,j) \in \mathsf{I} \times \mathsf{J}} = \{r_0^0(0|, s_0^0(0|, r_0^1(0|, s_0^1(0|, \ldots, r_j^i(\mathsf{j}|, s_j^i(\mathsf{j}|, \ldots\} \in \mathsf{Eff}(A),$$

where $r_j^i = \min\{p_j^i, q_j\}$ and $s_j^i = \max\{p_j^i, q_j^i\} - \min\{p_j^i, q_j^i\}$, one can verify from direct calculation that this is an effect of the theory, since it complies with Postulate 2, and that it satisfies the following relations,

$$(a_x|_A = \sum_{(i,j) \in \mathsf{V}_x} (c_{(i,j)}|_A \quad \forall x \in \mathsf{X},$$

$$(b_y|_A = \sum_{(i,j) \in \mathsf{W}_y} (c_{(i,j)}|_A \quad \forall y \in \mathsf{Y},$$

where the ensembles $\{\mathsf{V}_x\}_{x \in \mathsf{X}}$ and $\{\mathsf{W}_y\}_{y \in \mathsf{Y}}$ are appropriate disjoint partitions of $\mathsf{I} \times \mathsf{J}$. The proof is now concluded since we have shown that the two observation tests are compatible.

### 2. A property of MCT's tests

An interesting aspect of MCT's tests is that the ancillary system C can always be neglected and considered only through a coarse-graining operation.

To show this it is sufficient to observe that any state of the theory can be uniquely decomposed on the vertices of the simplex which is the state space, and that any effect can be written as coarse graining of the effects that perfectly discriminate the vertices of the simplex:

$$|\rho)_A = \sum_{i=1}^{\mathsf{D}_A} p_i|i)_A,$$

$$(a|_A = \sum_{j=1}^{\mathsf{D}_A} d_j(\mathsf{j}|_A,$$

where $p_i \in [0, 1] \forall i = 1, \ldots, \mathsf{D}_A$ and $\sum_{i=1}^{\mathsf{D}_A} \leqslant 1$, and $d_j \in [0, 1] \forall j = 1, \ldots, \mathsf{D}_A$.

This property still holds also for states and effects of composite systems. Therefore, any transformation of MCT obtained as parallel and sequential composition of the elements in (D2) can be rewritten as



$$= \sum_{i=1}^{D_{B'}} \sum_{j=1}^{D_{A'}} \sum_{m=1}^{D_C} \sum_{m'=1}^{D_C} p_{mi} d_{m'j}$$



$$= \sum_{i=1}^{D_{B'}} \sum_{j=1}^{D_{A'}} \sum_{m'=1}^{D_C} p_{mi} d_{m'j}$$



$$= \sum_{i=1}^{D_{B'}} \sum_{j=1}^{D_{A'}} \sum_{m=1}^{D_C} \sum_{m'=1}^{D_C} p_{mi} d_{m'j} \delta_{m,m'}$$



$$= \sum_{i=1}^{D_{B'}} \sum_{j=1}^{D_{A'}} \sum_{m=1}^{D_C} p_{mi} d_{mj}$$



.

[1] W. Heisenberg, Z. Phys. **43**, 172 (1927).

[2] H. P. Robertson, Phys. Rev. **34**, 163 (1929).

[3] P. J. Lahti, Int. J. Theor. Phys. **19**, 905 (1980).

[4] C. A. Fuchs and A. Peres, Phys. Rev. A **53**, 2038 (1996).

[5] M. Ozawa, Phys. Rev. A **67**, 042105 (2003).

[6] G. D'Ariano, Fortschr. Phys. **51**, 318 (2003).

[7] P. Busch, "No information without disturbance": Quantum limitations of measurement, in *Quantum Reality, Relativistic Causality, and Closing the Epistemic Circle: Essays in Honour of Abner Shimony* (Springer, Dordrecht, 2009), pp. 229–256.

[8] P. Busch, P. Lahti, and R. F. Werner, Rev. Mod. Phys. **86**, 1261 (2014).

[9] T. Heinosaari, Phys. Rev. A **93**, 042118 (2016).

[10] T. Heinosaari, L. Leppäjärvi, and M. Plávala, Quantum **3**, 157 (2019).

[11] From now on, we will use the word *observation* to denote a measurement where the output system is discarded, or simply disregarded. Notice that any measurement determines an observation, corresponding to the measurement itself followed by a discard of the output system.

[12] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).

[13] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Quantum from principles, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer, Dordrecht, 2016), pp. 171–221.

[14] G. M. D'Ariano, G. Chiribella, and P. Perinotti, *Quantum Theory from First Principles: An Informational Approach* (Cambridge University Press, Cambridge, U.K., 2017).

[15] G. M. D'Ariano, M. Erba, and P. Perinotti, Phys. Rev. A **102**, 052216 (2020).

[16] J. G. Richens, J. H. Selby, and S. W. Al-Safi, Phys. Rev. Lett. **119**, 080503 (2017).

[17] L. Catani, M. Leifer, G. Scala, D. Schmid, and R. W. Spekkens, Phys. Rev. Lett. **129**, 240401 (2022).

[18] R. W. Spekkens, Phys. Rev. A **75**, 032110 (2007).

[19] R. W. Spekkens, Quasi-quantization: Classical statistical theories with an epistemic restriction, in *Quantum Theory: Informational Foundations and Foils*, edited by G. Chiribella and R. W. Spekkens (Springer, Dordrecht, 2016), pp. 83–135.

[20] M. Plávala and M. Kleinmann, Phys. Rev. Lett. **128**, 040405 (2022).

[21] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, Phys. Rev. A **70**, 062101 (2004).

[22] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, Phys. Rev. A **86**, 012103 (2012).

[23] D. Schmid, J. H. Selby, E. Wolfe, R. Kunjwal, and R. W. Spekkens, PRX Quantum **2**, 010331 (2021).

[24] D. Schmid, J. H. Selby, M. F. Pusey, and R. W. Spekkens, A structure theorem for generalized-noncontextual ontological models, arXiv:2005.07161.

[25] G. M. D'Ariano, P. Perinotti, and A. Tosini, Quantum **4**, 363 (2020).

[26] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007).

[27] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, arXiv:quant-ph/0611295.

[28] G. Lüdwig, *An Axiomatic Basis for Quantum Mechanics* (Springer, Berlin, 1985).

[29] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens, Europhys. Lett. **103**, 10002 (2013).

[30] F. Buscemi, E. Chitambar, and W. Zhou, Phys. Rev. Lett. **124**, 120401 (2020).

[31] G. M. D'Ariano, P. Perinotti, and A. Tosini, J. Phys. A: Math. Theor. **55**, 394006 (2022).

[32] G. M. D'Ariano, M. Erba, and P. Perinotti, Phys. Rev. A **101**, 042118 (2020).

[33] More precisely, systems A and B are generally composite $A = A_1 A_2 \ldots A_j$, $B = B_1 B_2 \ldots B_k$, and $\mathcal{S}^{(1)}$ permutes the subsystems of A, while $\mathcal{S}^{(2)}$ permute those of system B.

[34] M. Plávala, Phys. Rev. A **94**, 042108 (2016).

[35] F. Buscemi, K. Kobayashi, S. Minagawa, P. Perinotti, and A. Tosini, Quantum **7**, 1035 (2023).

[36] We remark that the property of classical control on outcomes is actually quite strong, implying, e.g., the causality principle (see Sec. III and Ref. [14]).