Classical theories with entanglement

Giacomo Mauro D'Ariano,* Marco Erba,† and Paolo Perinotti‡
Università degli Studi di Pavia, Dipartimento di Fisica, QUIT Group, and INFN Gruppo IV, Sezione di Pavia, via Bassi 6, 27100 Pavia, Italy
(Dated: September 17, 2019)

We consider theories where the set of states for every system is a simplex. A special case of such theories is that of classical theories, i.e. those whose pure states are jointly perfectly discriminable. The standard classical theory is the one satisfying local discriminability. However, simplicial theories—including the classical ones—can violate local discriminability, thus admitting entangled states. After proving that simplicial theories are necessarily causal, we provide necessary and sufficient conditions for them to exhibit entanglement, and classify their system-composition rules.

Entanglement is the quantum feature marking the starkest departure of Quantum Theory from Classical Theory (the latter, in the following, will be referred to as *the standard classical theory*). The phenomenon is commonly popularized as the so-called *quantum nonlocality*, although the two concepts are not coincident. Indeed, it is known that the mere existence of entangled states is not sufficient for nonlocality [1, 2]. On the other hand, the standard classical theory does not allow for any kind of entanglement or nonlocality. In this letter, we argue that the absence of entanglement in the standard classical theory is due not only to the simplicial structure of the set of states of the single systems, but also to the composition rule of systems, which satisfies local discriminability of states. Indeed, as we will show in this letter, there exist *simplicial theories* that exhibit entanglement, and these include *classical theories*, defined as those simplicial theories where the pure states are jointly perfectly discriminable [3–5]. Besides classical theories, the definition of simplicial theory encompasses more general cases—e.g. noisy versions of classical theories, where pure states cannot be reliably distinguished due to a limited set of effects. The results of the present letter hold in the general case of simplicial theories. The rationale of such theories is that every deterministic state has a unique decomposition into pure states.

We prove that, among the classical theories, the only theory that does not admit entanglement is the standard classical one. We provide a classification of simplicial theories in terms of the composition rule for systems. Finally, as a consequence of the geometric structure of the sets of states, we show that a simplicial theory is necessarily causal, highlighting that causality is intrinsic to the standard classical theory. After characterising and classifying simplicial theories from the point of view of the discriminability and composition of systems, we draw our conclusions, discussing some information-theoretic features of this family of theories, and making a comparison with the existing literature.

We now briefly recall some notion of Operational Probabilistic Theories (OPTs) [6]. In the following we will denote by roman letters $\mathrm{A}, \mathrm{B}, \ldots \in \mathsf{Sys}(\Theta)$ the systems of the theory Θ, by $\mathscr{E} = \{\mathscr{E}_x\}_{x \in \mathrm{X}} \subseteq \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ a test of possible transformations $\mathscr{E}_x$ from input system A to output system B corresponding to outcomes x in the outcome space X. For transformations $\mathscr{A} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and $\mathscr{B} \in \mathsf{Transf}(\mathrm{B} \to \mathrm{C})$, we define their composition $\mathscr{B}\mathscr{A} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{C})$, and analogously for tests. We denote by $\mathsf{St}(\mathrm{A})$ and $\mathsf{Eff}(\mathrm{A})$ the set of states and effects of system A, respectively. We remind that states and effects are special cases of transformations—i.e. $\mathsf{St}(\mathrm{A}) = \mathsf{Transf}(\mathrm{I} \to \mathrm{A})$ and $\mathsf{Eff}(\mathrm{A}) = \mathsf{Transf}(\mathrm{A} \to \mathrm{I})$, where I is the trivial system—thus every result holding for transformations also applies to states and effects. We recall two important notions from convex and conic analysis, i.e. those of extremal and atomic points. Let C_+ be a convex cone, with $x_1, x_2 \in \mathsf{C}_+$ and $p \in (0,1)$. A point $x \in \mathsf{C}_+$ is called *atomic* if $x = x_1 + x_2$ implies $x_1 \propto x_2$, while it is called *extremal* if $x = px_1 + (1-p)x_2$ implies $x_1 = x_2$. We will denote by $\mathsf{ExtSt}(\mathrm{A})$ the set extremal points of the convex hull of $\mathsf{St}(\mathrm{A})$. Notice that $\mathsf{ExtSt}(\mathrm{A})$ contains the null "state" 0. The deterministic extremal states are historically called *pure states*.

We will also use the Dirac-like notation $|\rho)_\mathrm{A} \in \mathsf{St}(\mathrm{A})$ and $(a|_\mathrm{A} \in \mathsf{Eff}(\mathrm{A})$. Accordingly, given an arbitrary state $\rho_i \in \mathsf{St}(\mathrm{A})$, a transformation $\mathscr{E}_x \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$, and an effect $a_k \in \mathsf{St}(\mathrm{B})$, an Operation Probabilistic Theory (OPT) allows to compute joint probabilities of the form

$$p\,(i, x, k \mid \rho, \mathscr{E}, a) \coloneqq (a_k|_\mathrm{B} \mathscr{E}_x |\rho_i)_\mathrm{A}.$$

We finally remind the notion of *dimension* D_A of a system A, defined as $D_\mathrm{A} \coloneqq \dim \mathsf{St}_\mathbb{R}(\mathrm{A})$, where $\mathsf{St}_\mathbb{R}(\mathrm{A}) \coloneqq \mathsf{Span}_\mathbb{R} \mathsf{St}(\mathrm{A})$. Effects are *separating* for states (and viceversa), i.e. if two states $\rho_0, \rho_1 \in \mathsf{St}_\mathbb{R}(\mathrm{A})$ are different, there exists an effect $a \in \mathsf{Eff}_\mathbb{R}(\mathrm{A})$ such that $(a|\rho_0) \neq (a|\rho_1)$. The latter property is equivalent to the requirement that $\dim \mathsf{St}_\mathbb{R}(\mathrm{A}) = \dim \mathsf{Eff}_\mathbb{R}(\mathrm{A})$. As a standard requirement, we consider finite-dimensional OPTs, with $D_\mathrm{A} < +\infty$ for all $\mathrm{A} \in \mathsf{Sys}(\Theta)$. This means that systems of the theory can be completely probed via the statistics of a finite number of experiments. For instance, in the standard classical theory D_A is the number of perfectly distinguishable states of A, while in Quantum Theory $D_\mathrm{A} = d_\mathrm{A}^2$, where d_A is the dimension of the Hilbert space associated to system A.

The existence of parallel composition entails a prescription to assign a dimension D_{AB} to a composite system AB as a function of the dimensions D_{A} and D_{B} of the local systems A,B. $\mathsf{St}(\mathrm{AB})$ contains at least the product of the states of A and B, which can be composed independently, and similarly for effects. By virtue of the properties of the parallel composition, for every OPT Θ one has the inequality [7]:

$$D_{\mathrm{AB}} \geq D_{\mathrm{A}} D_{\mathrm{B}} \quad \forall \mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta). \tag{1}$$

This leads us to introduce the excess dimension $\Delta^{(2)}_{\mathrm{AB}} \geq 0$ of the composite system AB as follows

$$D_{\mathrm{AB}} =: D_{\mathrm{A}} D_{\mathrm{B}} + \Delta^{(2)}_{\mathrm{AB}} \quad \forall \mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta). \tag{2}$$

In both standard classical and quantum theory one has $\Delta^{(2)}_{\mathrm{AB}} = 0$. We now see how this relates to the degree of holism required in the task of state-discrimination.

Property 1 (n-local discriminability)**.** *For m-partite states and $n < m$, the factorized effects involving only k-partite effects with $k \leq n$ are separating.*

In the simplest case where $n = 1$, the property is called *local discriminability*. In both standard classical and quantum theory, the products of local effects are separating for multipartite states, namely both theories satisfy local discriminability.

Proposition 1. *Let Θ be an OPT. Then Θ satisfies local discriminability if and only if the following rule holds:*

$$D_{\mathrm{AB}} = D_{\mathrm{A}} D_{\mathrm{B}} \quad \forall \mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta). \tag{3}$$

Proof. The proof can be found in Ref. [8]. ∎

If a theory Θ does not satisfy local discriminability, then there exist $\mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta)$ such that $\Delta^{(2)}_{\mathrm{AB}} > 0$. Another important feature shared by both standard classical and quantum theory, but not by general theories, is the following.

Property 2 (atomicity of state-composition)**.** *The parallel composition of two atomic states is atomic [9].*

In this letter, we connect the properties of local discriminability and atomicity of state-composition to the presence of entangled states. We now recall the definition of separable and entangled states. Given two systems A,B, the *separable states* of the bipartite system AB are those of the form:

$$|\sigma)_{\mathrm{AB}} = \sum_{i \in I} p_i |\alpha_i)_{\mathrm{A}} |\beta_i)_{\mathrm{B}}, \tag{4}$$

with $p_i > 0$ for all $i \in I$, and $\sum_{i\in I} p_i \leq 1$. This class of states can be prepared using only local operations and classical communication. By negation, *entangled states* are those states that are *non-separable*. The two following results hold for arbitrary OPTs.

Proposition 2. *Let Θ be an OPT. If Θ does not satisfy local discriminability, then it admits entangled states.*

Proof. By hypothesis, from Eqs. (1) and (3) we have that there exist $\mathrm{A},\mathrm{B} \in \mathsf{Sys}(\Theta)$ such that $D_{\mathrm{AB}} > D_{\mathrm{A}} D_{\mathrm{B}}$. Since product states generate a subspace of $\mathsf{St}_{\mathbb{R}}(\mathrm{AB})$ of dimension $D_{\mathrm{A}} D_{\mathrm{B}}$, containing all separable states, $\mathsf{St}(\mathrm{AB})$ must contain at least a state that is not separable. ∎

Proposition 3. *Let Θ be an OPT. If Θ does not satisfy atomicity of state-composition, then it admits entangled states.*

Proof. By hypothesis, there exist two atomic states $|\rho)_{\mathrm{A}}, |\sigma)_{\mathrm{B}}$ whose product $|\pi)_{\mathrm{AB}} = |\rho)_{\mathrm{A}} |\sigma)_{\mathrm{B}}$ is not atomic. On the other hand, $|\pi)_{\mathrm{AB}}$ cannot be the form of Eq. (4), otherwise at least one state among $|\rho)_{\mathrm{A}}$ and $|\sigma)_{\mathrm{B}}$ would be not atomic, contradicting the hypothesis. This implies that there exists a state of AB which is not separable. ∎

In the following, we characterise simplicial theories from the point of view of the admissibility of entangled states, finding that the converse of Proposition 2 holds in general for this family of theories, whereas the converse of Proposition 3 holds in the case where a simplicial theory satisfies n-local discriminability for some positive integer n.

Definition 1 (simplicial theory)**.** *A simplicial theory Θ is a finite-dimensional OPT where the set of states of every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$ is a simplex in D_{A} dimensions.*

Since the sets of states are simplexes, in a simplicial theory the decomposition of every state into extremal states is unique.

Property 3 (causality)**.** *An OPT is causal if and only if it admits a unique deterministic effect [8].*

Theorem 1. *Simplicial theories are causal.*

Proof. Let Θ be a simplicial theory, $\mathrm{A} \in \mathsf{Sys}(\Theta)$ and $|i)_{\mathrm{A}}$ an extremal and non-null state. Suppose that $|i)_{\mathrm{A}}$ is not deterministic: then it can be completed to a deterministic one, namely there exists a state $|\sigma)_{\mathrm{A}}$ such that $|\rho)_{\mathrm{A}} = |i)_{\mathrm{A}} + |\sigma)_{\mathrm{A}}$ is deterministic. Now, let $(e_j|_{\mathrm{A}}$ for $j = 1, \ldots, D_{\mathrm{A}}$ be the linear functionals such that $(e_j|j') = \delta_{jj'}$ for every non-null extremal state $|j')_{\mathrm{A}}$: this implies that $0 \leq (e_j|\tau) \leq 1$ for all j and all $|\tau)_{\mathrm{A}} \in \mathsf{St}(\mathrm{A})$. One then has:

$$(e_i|\rho) = (e_i|i) + (e_i|\sigma) = 1 + (e_i|\sigma) \leq 1.$$

namely $(e_i|\sigma) = 0$ and $(e_i|\rho) = 1$. Then

$$|\sigma)_{\mathrm{A}} = \sum_{k=0}^{D_{\mathrm{A}}} p_k |k)_{\mathrm{A}}, \qquad p_k \geq 0\ \forall k,\ p_i = 0,\ \sum_{k=0}^{D_{\mathrm{A}}} p_k = 1.$$

Let us now pose

$$(e|_{\mathrm{A}} := \sum_{j=1}^{D_{\mathrm{A}}} (e_j|_{\mathrm{A}}.$$

Clearly $0 \leq (e|\tau) \leq 1$ for all $|\tau)_{\mathrm{A}} \in \mathsf{St}(\mathrm{A})$, and $(e|i) = 1$, implying $(e|\sigma) = 0$, i.e. $p_k = 0$ $\forall k \neq 0$. This shows that $|i)$ is deterministic. Being independent of i, the above argument proves that all non-null extremal states are deterministic. Now the effect $(e|_{\mathrm{A}}$ amounts to unit on all deterministic states, hence it is deterministic. Since the non-null extremal states of a simplicial theory are complete and linearly independent, there exists a unique effect $(e|_{\mathrm{A}}$ such that $(e|j)_{\mathrm{A}} = 1$ for all extremal non-null states $|j)_{\mathrm{A}}$. Thus the deterministic effect $(e|_{\mathrm{A}}$ is unique. ■

Our first result states that, for a simplicial theory, the converse of Proposition 2 is also true.

Theorem 2. *Let Θ be a simplicial theory. Then Θ admits entangled states if and only if it does not satisfy local discriminability.*

Proof. ($\Leftarrow$) The implication holds true by Proposition 2. ($\Rightarrow$) By hypothesis, the theory Θ is simplicial. In particular, this implies that, for every system $\mathrm{A} \in \mathsf{Sys}(\Theta)$: (i) A has exactly D_{A} non-null extremal states, and (ii) every deterministic state of A has a unique decomposition as a convex combination of non-null extremal states. Let us denote the non-null extremal states of a system X by $|l)_{\mathrm{X}}$, for $l = 1, \ldots, D_{\mathrm{X}}$. By Theorem 1 such states are also deterministic. Then, for all $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and all $|i)_{\mathrm{A}} \in \mathsf{ExtSt}(\mathrm{A}), |j)_{\mathrm{B}} \in \mathsf{ExtSt}(\mathrm{B})$, there exists a non-empty set $I_{ij} \subseteq \{1, \ldots, D_{\mathrm{AB}}\}$ such that

$$|i)_{\mathrm{A}}|j)_{\mathrm{B}} = \sum_{k \in I_{ij}} p_k^{ij} |k)_{\mathrm{AB}}, \tag{5}$$

with $p_k^{ij} > 0$ for $k \in I_{ij}$ and $\sum_{k \in I_{ij}} p_k^{ij} = 1$. Let $(e|_{\mathrm{A}} \in \mathsf{Eff}(\mathrm{A})$ be the (unique) deterministic effect of A. Moreover, let us suppose that there exist two different pairs of indices $(\tilde{i}, \tilde{j}), (\tilde{i}', \tilde{j}')$ such that $\tilde{J} := I_{\tilde{i}\tilde{j}} \cap I_{\tilde{i}'\tilde{j}'} \neq \emptyset$. Without loss of generality, we can assume $\tilde{j} \neq \tilde{j}'$. Choose $\tilde{k} \in \tilde{J}$: since for all $k \in \{1, \ldots, D_{\mathrm{A}}D_{\mathrm{B}}\}$ $(e|_{\mathrm{A}}|k)_{\mathrm{AB}} \in \mathsf{St}(\mathrm{B})$ is deterministic, it is also non-vanishing for all k, and being the decomposition into non-null extremal states unique by simpliciality, from Eq. (5) we conclude that $(e|_{\mathrm{A}}|\tilde{k})_{\mathrm{AB}} = |\tilde{j})_{\mathrm{B}} = |\tilde{j}')_{\mathrm{B}}$, which is absurd. Thus $I_{ij} \cap I_{i'j'} = \emptyset$ for every different pair of indices. Now, let us suppose that Θ satisfies local discriminability. Then, by Proposition 1, Eq. (3) holds, and $I_{ij} \subseteq \{1, 2, \ldots, D_{\mathrm{A}}D_{\mathrm{B}}\}, \forall i, j$. By conditions $I_{ij} \neq \emptyset$ and $I_{ij} \cap I_{i'j'} = \emptyset$, every set I_{ij} must be a singleton. This implies that for every pair i, j there exists $|k)_{\mathrm{AB}}$ such that $|k)_{\mathrm{AB}} = |i)_{\mathrm{A}}|j)_{\mathrm{B}}$, namely the parallel composition of non-null extremal states is a non-null extremal state. Accordingly, Eq. (3) implies that the states of the form $|i)_{\mathrm{A}}|j)_{\mathrm{B}}$ exhaust the set of non-null extremal states of the composite system AB. As a consequence, Θ does not admit entangled states. ■

Our second result states that, for a simplicial theory satisfying n-local discriminability for some positive integer n, the converse of Proposition 3 is also true. For this we need the following result, that is proved in the Supplemental Material.

Proposition 4. *Let Θ be a simplicial OPT satisfying n-local discriminability for some positive integer n. Suppose that there exist a pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and an extremal state $|\lambda)_{\mathrm{AB}} \in \mathsf{ExtSt}(\mathrm{AB})$ that does not convexly refine any product state $|\rho)_{\mathrm{A}}|\sigma)_{\mathrm{B}} \in \mathsf{St}(\mathrm{AB})$. Then $|\lambda)_{\mathrm{AB}} = 0$.*

We are now in position to prove the following theorem.

Theorem 3. *Let Θ be a simplicial theory satisfying n-local discriminability for some positive integer n. Then Θ admits entangled states if and only if it does not satisfy atomicity of state-composition.*

Proof. ($\Leftarrow$) The implication holds true by Proposition 3. ($\Rightarrow$) According to Proposition 4, under the hypothesis of the theorem every non-null extremal state of a bipartite system refines some product of extremal states. Since by hypothesis the theory admits entangled states, by Theorem 2 and Proposition 1 there exists a pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ such that $D_{\mathrm{AB}} > D_{\mathrm{A}}D_{\mathrm{B}}$. Then, there are more non-null extremal states of the composite system than products of (non-null) extremal states. Thus, there must be a product of extremal states $|i)_{\mathrm{A}}|j)_{\mathrm{B}}$ that is refined by more than one non-null extremal states, i.e. $|i)_{\mathrm{A}}|j)_{\mathrm{B}} = \sum_{k \in I_{ij}} p_k^{ij} |k)_{\mathrm{AB}}$, namely atomicity of state-composition does not hold. ■

The following corollary provides a characterisation of simplicial theories with entanglement. The standard classical theory is the unique classical theory where local discriminability holds.

Corollary 1. *Let Θ be a simplicial theory satisfying n-local discriminability for some positive integer n. Then, for all systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and non-null extremal states $|k)_{\mathrm{AB}} \in \mathsf{ExtSt}(\mathrm{AB})$, there exists a unique product of non-null extremal states $|i_k j_k)_{\mathrm{AB}} = |i_k)_{\mathrm{A}}|j_k)_{\mathrm{B}}$ such that $|k)_{\mathrm{AB}}$ convexly refines $|i_k j_k)_{\mathrm{AB}}$.*

Proof. Existence is provided by Proposition 4 and uniqueness has been proved in the proof of Theorem 2. ■

The latest result leads to a classification of the classical theories from the point of view of their degree of nonlocality. Notice that the map $k \mapsto i_k j_k$ used in the above proof is not injective in general. As a straightforward consequence of the previous corollary, we are now in position to provide a general classification of the composite state-spaces in simplicial theories satisfying n-local discriminability for some integer n.

Theorem 4 (classification of composite state-spaces in simplicial theories)**.** *Let Θ be a simplicial theory satisfying n-local discriminability for some integer n. For every pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, the state-space $\mathsf{St}\,(\mathrm{AB})$ is classified by the following: (i) the dimension of the composite system D_{AB}; (ii) the pure states of AB can be labelled as $|(ij)_k)_{\mathrm{AB}}$, for $i \in \{1, \ldots, D_{\mathrm{A}}\}$, $j \in \{1, \ldots, D_{\mathrm{B}}\}$, and k in a finite set I_{ij}, such that:*

$$|i)_{\mathrm{A}}|j)_{\mathrm{B}} = \sum_{k \in I_{ij}} p_k^{ij} |(ij)_k)_{\mathrm{AB}} \tag{6}$$

with $p_k^{ij} > 0$ and $\sum_{k \in I_{ij}} p_k^{ij} = 1$ for all non-null extremal states $|i)_{\mathrm{A}} \in \mathsf{ExtSt}(\mathrm{A}), |j)_{\mathrm{B}} \in \mathsf{ExtSt}(\mathrm{B})$ [10].

Discussion

We showed that every simplicial theory which is not the standard classical theory contains entangled states. We also characterised simplicial theories with entanglement, proving that they cannot satisfy local discriminability or—under the hypothesis of n-local discriminability for an integer $n > 1$—cannot satisfy atomicity of state-composition. This entails that there exists a pair of systems A, B such that at least a product of pure states of A and B is not pure. Finally, we proved that every simplicial theory is causal. The results of the present letter do not rely on any additional structure—such as perfect discriminability, or other principles—beyond the simplicial one.

If a theory is simplicial, there is no complementarity, and it is impossible to violate Bell's Inequalities [11]. Moreover, in presence of local discriminability, the converse is also true [12]. Nevertheless, all simplicial theories with entanglement contain states with non-null discord (see Ref. [13]), despite being simplicial. Furthermore, it is interesting to notice that simplicial theories have no *dimension mismatch* [14] and satisfy the *Information Content Principle* [15]. Indeed, these properties are not sensitive to the parallel composition rule, making all the simplicial theories equivalent on the basis of their simplicial structure. However, the simplicial theories with entanglement remarkably feature *hypersignalling* [16], thus displaying time-like correlations which are stronger than those of both Classical and Quantum Theory. At this stage, it is not possible to compare simplicial theories that exhibits entanglement with works on *broadcasting* [4] or *teleportation* [17] in generalized probabilistic theories, since local discriminability was therein assumed.

One could then examine the relation between simplicial theories with entanglement and theories having an epistemic restriction—such as Spekkens' toy theory [18]—or consider them in the light of Refs. [19, 20] about emergent classicality. An explicit construction of a simplicial theory with entanglement, complete with the set of transformations, would allow one to study the simplicial scenario as far as information causality [21] or communication complexity are concerned.

Finally, as shown in Ref. [22], simplicial theories are the only ones where full information can be extracted by measurements that do not disturb the measured system (and this in particular must hold even in the presence of entanglement).

SUPPLEMENTAL MATERIAL

Operational Probabilistic Theories

The primitive notions of an operational theory are those of *systems*, *tests*, and *events*. A system S represents the physical entity which is probed in the laboratory, such as a radiation field, a molecule, an elementary particle. A test $\mathscr{E}$ is characterized by a collection of events $\{\mathscr{E}_x\}_{x \in \mathrm{X}}$, and represents the single occurrence of a physical process, such as the use of a physical device or a measuring apparatus. The *outcome space* X associated with a test $\mathscr{E}$ collects all the possible outcomes that can occur within the test. Each test $\mathscr{E}$ is characterised by an input system A and an output system B. For instance, think of an electron–proton scattering: both the input and the output systems are an electron–proton pair, the test contains only one event corresponding to the two-particle interaction, and finally the outcome space is the singleton set. Whenever the output of a test $\mathscr{E}_1$ coincides with the input of another test $\mathscr{E}_2$, the *sequential composition* $\mathscr{E}_2 \circ \mathscr{E}_1 := \mathscr{E}_2\mathscr{E}_1$ of the two tests can be defined, being an allowed test for the theory. Thus, physical systems define the connection rules for tests.

In an Operational Probabilistic Theory (OPT), two important classes of tests are those of *preparations* and *observations*, i.e. tests where there is no input or no output system, respectively. Preparations ρ and observations a, where A is the output or the input system, are denoted using the Dirac-like notation $|\rho)_{\mathrm{A}}$ and $(a|_{\mathrm{A}}$. Given some arbitrary preparation event $|\rho_i)_{\mathrm{A}}$, event $\mathscr{E}_x$ from system A to system B respectively, and observation event $(a_k|_{\mathrm{B}}$, the purpose of an OPT is to compute joint probabilities of the form:

$$p\,(i, x, k \mid \rho, \mathscr{E}, a) := (a_k|_{\mathrm{B}}\mathscr{E}_x|\rho_i)_{\mathrm{A}}. \tag{7}$$

Two events from system X to system Y are equivalent if all their joint probabilities of the form (7), for the same other events appearing in Eq. (7), are equal. We will call equivalent classes of preparation and observation events *states* and *effects*, respectively. Besides, equivalent classes of arbitrary events from A to B will be called *transformations*. Given an OPT Θ, $\mathsf{Sys}(\Theta)$ will denote the set of the systems of the theory. For every $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$, $\mathsf{St}(\mathrm{A})$, $\mathsf{Transf}(\mathrm{A} \to \mathrm{B})$, and $\mathsf{Eff}(\mathrm{B})$ will denote, respectively, the sets of states of A, of transformations from A to B, and of effects of B.

Clearly, every event $\mathscr{E}_x \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ is a map from $\mathsf{St}(\mathrm{A})$ to $\mathsf{St}(\mathrm{B})$, or, dually, from $\mathsf{Eff}(\mathrm{B})$ to $\mathsf{Eff}(\mathrm{A})$. According to the above definitions, effects are separating for states, i.e. given two states $|\rho_1)_\mathrm{A} \neq |\rho_2)_\mathrm{A}$, there exists $(a|_\mathrm{A} \in \mathsf{Eff}(\mathrm{A})$ such that $(a|\rho_1) \neq (a|\rho_2)$. Similarly, an effect is the equivalence class of those observation events that give the same probabilities for every state, and thus states are separating for effects.

Within an operational perspective, an agent is also allowed to perform a test—say with outcome space X—disregarding the outcomes of a subset $Y \subseteq X$, and then merging events in Y into a single event: this possibility is captured by the notion of *coarse-graining*. According to probability theory, the probability of the coarse-grained event Y amounts to the sum of the probabilities of all the outcomes in the subset Y. Then, for each test $\{\mathscr{E}_x\}_{x \in X} \subseteq \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and every subset $Y \subseteq X$, the coarse-grained event is symbolically given by $\sum_{y \in Y} \mathscr{E}_y$, and sequential composition distributes over sums. The converse procedure of a coarse-graining is called a *refinement*. An event with trivial refinement, i.e. which cannot be further refined modulo a rescaling by a probability, is called *atomic*. Being the framework probabilistic, one may also consider convex combinations of states, transformations, and effects, corresponding to a randomization, i.e. a statistical mixture of events. In general, states, transformations, and effects can be thought of as embedded in convex spaces. It is often convenient to consider the real span of sets of states $\mathsf{St}(\mathrm{A})$, which is a linear space denoted by $\mathsf{St}_\mathbb{R}(\mathrm{A}) := \mathsf{Span}_\mathbb{R}\mathsf{St}(\mathrm{A})$. Every system A is then associated to a quantity $D_\mathrm{A} := \dim \mathsf{St}_\mathbb{R}(\mathrm{A})$, which is called *the dimension of the system* A. As a standard requirement, we consider finite-dimensional OPTs, namely theories where $D_\mathrm{A} < +\infty$ for all systems $\mathrm{A} \in \mathsf{Sys}(\Theta)$. The latter assumption means that we are considering systems that can be completely probed via the statistics of a finite number of experiments. For instance, in Classical Theory D_A is the number of perfectly distinguishable states of a system A, while in Quantum Theory $D_\mathrm{A} = d_\mathrm{A}^2$, where d_A is the dimension of the Hilbert space associated to the system A.

A test is called *deterministic* if the associated outcome space is the singleton set, i.e. the test has a single outcome. The interpretation of a deterministic test is that the physical process considered happens with certainty, i.e. with probability 1. For instance, a states is deterministic if and only if it gives probability 1 on every deterministic effect, or, in other words, if and only if it is normalized.

We conclude this review considering a fundamental structure of operational theories: *parallel composition*. Indeed, the last piece of information one needs to characterise an OPT is a recipe to form compounds out of systems and events available to local experimenters, i.e. which are uncorrelated. We denote parallel composition by the symbol $\boxtimes$. The main property of parallel composition is that

$$(\mathscr{A} \boxtimes \mathscr{B}) \circ (\mathscr{C} \boxtimes \mathscr{D}) = (\mathscr{A} \circ \mathscr{C}) \boxtimes (\mathscr{B} \circ \mathscr{D}).$$

In the case of systems, states, and effects, we will use the following notation: $\mathrm{AB} := \mathrm{A} \boxtimes \mathrm{B}$, $|\rho)_\mathrm{A}|\sigma)_\mathrm{B} := |\rho)_\mathrm{A} \boxtimes |\sigma)_\mathrm{B}$, and $(a|_\mathrm{A}(b|_\mathrm{B} := (a|_\mathrm{A} \boxtimes (b|_\mathrm{B}$. One has then to specify, for all events $\mathscr{E} \in \mathsf{Transf}(\mathrm{A} \to \mathrm{B})$ and $\mathscr{D} \in \mathsf{Transf}(\mathrm{C} \to \mathrm{D})$, how the composite event $\mathscr{E} \boxtimes \mathscr{D}$ embeds into the total space of events $\mathsf{Transf}(\mathrm{AC} \to \mathrm{BD})$. In Quantum Theory, for instance, this operation is given by the standard tensor product $\otimes$. As for deterministic events, the probabilistic structure implies that the parallel composition of two deterministic events is deterministic.

Proof of Proposition 4

We recall that in the main text of the letter we proved that: (i) any simplicial theory is causal (see Theorem 1), and (ii) the non-null extremal states of a simplicial theory are deterministic (see proof of Theorem 1). Accordingly, we will make use of the above results in the present section.

We also introduce the definition of separable state, which is of crucial relevance in the remainder. Let $\mathrm{S} = \mathrm{S_1 S_2} \cdots \mathrm{S_n} \in \mathsf{Sys}(\Theta)$, and $|\rho) \in \mathsf{St}(\mathrm{S})$. We say that $|\rho)$ is *separable* if there exist (finitely many) disjoint non-trivial bipartitions $\mathrm{S}_0^a := \{i_1^a, i_2^a, \dots, \imath_k^a\}$, $\mathrm{S}_1^a := \{j_1^a, j_2^a, \dots, j_{N-k}^a\}$ of $\{1, 2, \dots, n\}$ such that

$$|\rho) = \sum_{a \in \mathsf{A}} p_a |\sigma)_{\mathrm{S}_0^a} |\tau)_{\mathrm{S}_1^a},$$

with $p_a > 0$ for all $a \in \mathsf{A}$.

Lemma 1. *Let Θ be a simplicial OPT satisfying n-local discriminability for some positive integer n. For all $(n+1)$-partite system $\mathrm{S} = \mathrm{S_1 S_2} \cdots \mathrm{S_{n+1}} \in \mathsf{Sys}(\Theta)$, every state $|\rho) \in \mathsf{St}(\mathrm{S_1 S_2} \cdots \mathrm{S_{n+1}})$ admits a convex decomposition into states each of which convexly refines some separable state of $\mathrm{S_1 S_2} \cdots \mathrm{S_{n+1}}$.*

Proof. Take the subset $\mathcal{E} \subseteq \mathsf{ExtSt}(\mathrm{S_1} \cdots \mathrm{S_{n+1}})$ of all non-null extremal states of $\mathsf{St}(\mathrm{S_1} \cdots \mathrm{S_{n+1}})$ which convexly refine some separable state. Since Θ satisfies n-local discriminability, this is a spanning set for the space $\mathsf{St}_\mathbb{R}(\mathrm{S_1} \cdots \mathrm{S_{n+1}})$. Moreover, since Θ is simplicial, the elements of $\mathcal{E}$ are linearly independent. As a consequence, the dimension of $\mathsf{St}_\mathbb{R}(\mathrm{S_1} \cdots \mathrm{S_{n+1}})$ amounts to the cardinality of $\mathcal{E}$, which is then, by simpliciality, a complete set of states convexly generating every state $|\rho) \in \mathsf{St}(\mathrm{S_1 \cdots S_{n+1}})$. Equivalently, $\mathsf{ExtSt}(\mathrm{S_1 \cdots S_{n+1}}) = \{0\} \cup \mathcal{E}$. ■

Lemma 2. *Let Θ be a simplicial OPT. Let $|\pi) \in \mathsf{St}(\mathrm{S_1 \cdots S_n})$ with $n \geq 2$, so that $|\pi) = |\pi_I)|\pi_J)$, for some states $|\pi_I) \in \mathrm{S_I}, |\pi_J) \in \mathrm{S_J}$ with $I \cup J = \{1, \dots, n\}$, $I, J \neq \emptyset$, $I \cap J = \emptyset$, and $\mathrm{S_K} = \mathrm{S_{k_1} S_{k_2} \cdots S_{k_l}}$ for every l-tuple $K \subseteq \{1, \dots, n\}$. Let $|\phi) \in \mathsf{ExtSt}(\mathrm{S_1 \cdots S_n})$ be a non-null extremal state that convexly refines $|\pi)$. Finally, given $i \in I$ and $j \in J$, let $(e_{K_{ij}}|$ denote the deterministic effect on $\mathrm{S_{K_{ij}}}$ with $K_{ij} = \{1, \dots, n\} \setminus \{i, j\}$. Then $(e_{K_{ij}}|\pi)$ is a product state, and $(e_{K_{ij}}|\phi) \in \mathsf{St}(\mathrm{S_i S_j})$ is a physical state that convexly refines $(e_{K_{ij}}|\pi)$.*

Proof. The case $n = 2$ is trivially true, and we will then assume $n \geq 3$ in the following. By hypothesis, we can pose $|\pi) = p|\phi) + (1-p)|\sigma)$, where $|\sigma)$ is a deterministic state and $p \in (0, 1]$. By construction we have that

$$(e_{K_{ij}}|\pi) = p(e_{K_{ij}}|\phi) + (1-p)(e_{K_{ij}}|\sigma) \in \mathsf{St}(\mathrm{S_i S_j}), \quad (8)$$

and clearly both $(e_{K_{ij}}|\phi)$ and $(e_{K_{ij}}|\sigma)$ are deterministic states of $\mathrm{S_i S_j}$. Moreover, by causality, $(e_{K_{ij}}| = (e_{I \setminus \{i\}}|(e_{J \setminus \{j\}}|$, then $(e_{K_{ij}}|\pi)$ is a product state of $\mathrm{S_i S_j}$. Since the decomposition into non-null extremal states is unique by simpliciality, from Eq. (8) we conclude that $(e_{K_{ij}}|\phi)$—that is non-null, although it may possibly be non-extremal—convexly refines $(e_{K_{ij}}|\pi)$. ■

Proposition 8. *Let Θ be a simplicial OPT satisfying n-local discriminability for some positive integer n. Suppose that there exist a pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and an extremal state $|\lambda)_{\mathrm{AB}} \in \mathsf{ExtSt}(\mathrm{AB})$ that does not convexly refine any product state $|\rho)_{\mathrm{A}}|\sigma)_{\mathrm{B}} \in \mathsf{St}(\mathrm{AB})$. Then $|\lambda)_{\mathrm{AB}} = 0$.*

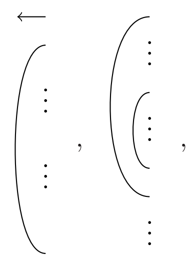
Proof. By hypothesis, there exist a pair of systems $\mathrm{A}, \mathrm{B} \in \mathsf{Sys}(\Theta)$ and an extremal state $|\lambda)_{\mathrm{AB}} \in \mathsf{ExtSt}(\mathrm{AB})$ that does not convexly refine any product state $|\rho)_{\mathrm{A}}|\sigma)_{\mathrm{B}} \in \mathsf{St}(\mathrm{AB})$. Let us denote

$|\lambda)_{\mathrm{AB}} \coloneqq$  . (9)

The theory Θ satisfies n-local discriminability. The case $n = 1$ is trivial, since all states are separable (see Theorem 2 in the main text of the letter) and $|\lambda)_{\mathrm{AB}}$ must be vanishing. In the following we will then assume $n \geq 2$. Let us now define the following $(n+1)$-partite state:

$|\Psi) \coloneqq$  $\in \mathsf{St}((\mathrm{A_1 \cdots A_n})\, \mathrm{B_1 \cdots B_n})$, (10)

where the systems A_m and $\mathrm{B}_{m'}$ are copies of, respectively, A and B for every $m, m' \in \{1, \dots, n\}$. By Lemma 1, the state $|\Psi)$ must be in the convex hull of some states refining the separable states $|\Lambda)$ of $(\mathrm{A_1 \cdots A_n})\, \mathrm{B_1 \cdots B_n}$. The latter must be of one of the following two types

$|\Lambda) =$  (11)

where in the first case $(\mathrm{A_1 \cdots A_n})$ is factorized from $\mathrm{B_1 \cdots B_n}$, while in the second case there must exists a state of some proper subsystem of $\mathrm{B_1 \cdots B_n}$ that is factorized. Then there exist coefficients $\alpha_i \in [0, 1]$ such that:

$$|\Psi) = \sum_i \alpha_i |\phi_i), \quad (12)$$

where the $|\phi_i)$ are non-null extremal states in the convex refinement of some separable state $|\Lambda)$ of one of the two types in Eq. (11). By construction, in both cases we can always find at least a subsystem $\mathrm{S} = \mathrm{A}_j \mathrm{B}_j$ of $\mathrm{A_1 \cdots A_n B_1 \cdots B_n}$ (now considered as a $2n$-partite system) such that the marginal state $(e_{\bar{\mathrm{S}}}|\Lambda)$ (where $\bar{\mathrm{S}}$ is the complementary subsystem of S in $\mathrm{A_1 \cdots A_n B_1 \cdots B_n}$) is a product state of $\mathrm{A}_j \mathrm{B}_j$. On the other hand the marginal state of $|\Psi)$ on S is $(e_{\bar{\mathrm{S}}}|\Psi) = |\lambda)_{\mathrm{A}_j \mathrm{B}_j}$. For each term on the r.h.s. in Eq. (12), one can then find the above mentioned subsystem S and apply $(e_{\bar{\mathrm{S}}}|$ to both sides. This gives an equation of the form:

$$|\lambda)_{\mathrm{A}_j \mathrm{B}_j} = \alpha_i |\chi_i)_{\mathrm{A}_j \mathrm{B}_j} + |\omega_i)_{\mathrm{A}_j \mathrm{B}_j}, \quad (13)$$

where $|\chi_i)_{\mathrm{A}_j \mathrm{B}_j}$ is in the convex refinement of some product state by Lemma 2, and $|\omega_i)_{\mathrm{A}_j \mathrm{B}_j}$ is a physical state. Since $|\lambda)_{\mathrm{A}_j \mathrm{B}_j}$ is an extremal point of a simplex, its convex decomposition must be trivial, and then either $\alpha_i = 0$ or

$$|\lambda)_{\mathrm{A}_j \mathrm{B}_j} = |\chi_i)_{\mathrm{A}_j \mathrm{B}_j}. \quad (14)$$

Finally, either $\alpha_i = 0$ holds for every i, and then by direct inspection of the definitions (9) and (10) $|\lambda)_{\mathrm{AB}}$ is vanishing, or identity (14) holds for some i, contradicting the hypothesis. ■

---

* dariano@unipv.it
† marco.erba@unipv.it
‡ paolo.perinotti@unipv.it

[1] Reinhard F. Werner, "Quantum states with Einstein-Podolsky-Rosen correlations admitting a hidden-variable model," Phys. Rev. A **40**, 4277–4281 (1989).

[2] Jonathan Barrett, "Nonsequential positive-operator-valued measurements on entangled mixed states do not always violate a Bell inequality," Phys. Rev. A **65**, 042302 (2002).
[3] Jonathan Barrett, "Information processing in generalized probabilistic theories," Phys. Rev. A **75**, 032304 (2007).
[4] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce, "Generalized No-Broadcasting Theorem," Phys. Rev. Lett. **99**, 240501 (2007).
[5] Corsin Pfister and Stephanie Wehner, "An information-theoretic principle implies that any discrete physical theory is classical," Nature Communications **4**, 1851 EP – (2013).
[6] For a synthetic review of OPTs one can see the supporting material to the present letter. For a thorough presentation see Refs. [8, 23–25].
[7] Lucien Hardy and William K. Wootters, "Limited Holism and Real-Vector-Space Quantum Theory," Foundations of Physics **42**, 454–473 (2012).
[8] Giacomo Mauro D'Ariano, Giulio Chiribella, and Paolo Perinotti, *Quantum Theory from First Principles: An Informational Approach* (Cambridge University Press, 2017).
[9] One may argue that the property of atomicity of state-composition should be restated as purity of state-composition. However, differently from the notion of *atomicity* (from conic analysis) and *extremality* (from convex analysis) the notion of *purity* is an historical one introduced only for extremal deterministic states. In different situations it is often used as a synonym of atomicity.
[10] Notice that not all choices of p_k^{ij} necessarily lead to a consistent theory. However, there are cases of classical theories beyond the standard ones, proving that a consistent choice of p_k^{ij} is possible [26].
[11] Indeed, if there is no complementarity, there exists a joint probability distribution for the outcomes of any pair of measurements, which then excludes a violation of any probability bound.
[12] Martin Plávala, "All measurements in a probabilistic theory are compatible if and only if the state space is a simplex," Phys. Rev. A **94**, 042108 (2016).
[13] Paolo Perinotti, "Discord and Nonclassicality in Probabilistic Theories," Phys. Rev. Lett. **108**, 120502 (2012).
[14] Nicolas Brunner, Marc Kaplan, Anthony Leverrier, and Paul Skrzypczyk, "Dimension of physical systems, information processing, and thermodynamics," New Journal of Physics **16**, 123050 (2014).
[15] L. Czekaj, M. Horodecki, P. Horodecki, and R. Horodecki, "Information content of systems as a physical principle," Phys. Rev. A **95**, 022119 (2017).
[16] Michele Dall'Arno, Sarah Brandsen, Alessandro Tosini, Francesco Buscemi, and Vlatko Vedral, "No-Hypersignaling Principle," Phys. Rev. Lett. **119**, 020401 (2017).
[17] Howard Barnum, Jonathan Barrett, Matthew Leifer, and Alexander Wilce, "Teleportation in general probabilistic theories," in *Proceedings of Symposia in Applied Mathematics*, Vol. 71 (2012) pp. 25–48.
[18] Robert W. Spekkens, "Evidence for the epistemic view of quantum states: A toy theory," Phys. Rev. A **75**, 032110 (2007).
[19] Miguel Navascués and Harald Wunderlich, "A glance beyond the quantum model," Proceedings of the Royal Society A: Mathematical, Physical and Eng
[20] Jonathan G. Richens, John H. Selby, and Sabri W. Al-Safi, "Entanglement is necessary for emergent classicality in all physical theories," Phys. Rev. Lett. **119**, 080503 (2017).
[21] Marcin Pawłowski, Tomasz Paterek, Dagomir Kaszlikowski, Valerio Scarani, Andreas Winter, and Marek Żukowski, "Information causality as a physical principle," Nature **461**, 1101 EP – (2009).
[22] Giacomo Mauro D'Ariano, Paolo Perinotti, and Alessandro Tosini, "Information and disturbance in operational probabilistic theories," arxiv.org (2019), 1907.07043.
[23] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, "Probabilistic theories with purification," Phys. Rev. A **81**, 062348 (2010).
[24] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, "Informational derivation of quantum theory," Phys. Rev. A **84**, 012311 (2011).
[25] Giulio Chiribella, Giacomo Mauro D'Ariano, and Paolo Perinotti, "Quantum from principles," in *Quantum Theory: Informational Foundations and Foils*, edited by Giulio Chiribella and Robert W. Spekkens (Springer Netherlands, Dordrecht, 2016) pp. 171–221.
[26] G. M. D'Ariano, M Erba, and P. Perinotti, Unpublished.