Quantum Inf Process (2010) 9:95–141
DOI 10.1007/s11128-010-0172-3

Testing axioms for quantum theory on probabilistic toy-theories

Giacomo Mauro D'Ariano · Alessandro Tosini

Received: 27 November 2009 / Accepted: 1 March 2010 / Published online: 19 March 2010

Abstract In D'Ariano in Philosophy of Quantum Information and Entanglement, Cambridge University Press, Cambridge, UK (2010), one of the authors proposed a set of operational postulates to be considered for axiomatizing Quantum Theory. The underlying idea is to derive Quantum Theory as the mathematical representation of a *fair operational framework*, i.e. a set of rules which allows the experimenter to make predictions on future *events* on the basis of suitable *tests*, e.g. without interference from uncontrollable sources and having local control and low experimental complexity. In addition to causality, two main postulates have been considered: PFAITH (existence of a pure preparationally faithful state), and FAITHE (existence of a faithful effect). These postulates have exhibited an unexpected theoretical power, excluding all known nonquantum probabilistic theories. In the same paper also postulate PURIFY-1 (purifiability of all states) has been introduced, which later has been reconsidered in the stronger version PURIFY-2 (purifiability of all states unique up to reversible channels on the purifying system) in Chiribella et al. (Reversible realization of physical processes in probabilistic theories, arXiv:0908.1583). There, it has been shown that Postulate PURIFY-2, along with causality and local discriminability, narrow the probabilistic theory to something very close to the quantum one. In the present paper we test the above postulates on some nonquantum probabilistic models. The first model—*the two-box world*—is an extension of the Popescu–Rohrlich model (Found Phys, 24:379, 1994), which achieves the greatest violation of the CHSH inequality compatible with the no-signaling principle. The second model—*the two-clock world*— is actually a full class of models, all having a disk as convex set of states for the local system.

G. M. D'Ariano (✉) · A. Tosini
Universita' di Pavia, Pavia, Italy
e-mail: dariano@unipv.it

One of them corresponds to—*the two-rebit world*— namely qubits with real Hilbert space. The third model—*the spin-factor*—is a sort of n-dimensional generalization of the clock. Finally the last model is *the classical probabilistic theory*. We see how each model violates some of the proposed postulates, when and how teleportation can be achieved, and we analyze other interesting connections between these postulate violations, along with deep relations between the local and the non-local structures of the probabilistic theory.

Keywords Axiomatization of quantum theory · Operational probabilistic theories · Toy-theories

1 Introduction

Quantum Theory is still laking a foundation. The Lorentz transformations suffered the same problem before the discovery of special relativity, and an analogous principle of "quantumness" has not been found yet. If one considers the theoretical power of special relativity in the ensuing research, one definitely ought to put the principle of quantumness at the highest research priority. Despite of it the research efforts in this direction have been sporadic.

A promising attack to the problem has been that of *quantum logic*, introduced in the early works of Birkoff, von Neumann, Jordan, and Wigner [1,2]. This approach, which proposes to regard Quantum Theory as a new kind of *propositional calculus* based on a logic different from the classical one, was neglected for a long time until the works of Varadarajan [3], and most notably by Mackey [4], who axiomatized Quantum Theory within an operational framework, with the single exception of an admittedly ad hoc postulate without physical significance. Another deep effort in justifying Quantum Theory came from Segal [5] who proposed a purely algebraic formulation instead of the usual Hilbert-space axiomatization, however leading to a much more general mathematical framework than the quantum one. Beside the cited logical and algebraic approaches we can mention the notable attempt of operational axiomatization by Ludwig and his school [6]. More recently in Ref. [7], has provided "five reasonable axioms" to derive Quantum Theory, which, however, are still not operational, e.g. the existence of a fixed function connecting the number of perfectly discriminable states of a system with the dimension of its convex set of states. However, many ideas introduced in this paper can offer a good starting point in the search of an operational axiomatization. Among other recent attempts of operational axiomatization of Quantum Theory, the one of Clifton, Bub, and Halvorson [8] gained some popularity, however, based on the common misunderstanding that the C*-algebraic formulation of observables is itself operational, whereas it is exactly equivalent to the Hilbert-space formulation, apart from the natural inclusion of super-selection rules, whence with the possibility of treating the classical and quantum cases contextually. On the basis of three fundamental information-theoretic constraints—(a) the no-signaling, (b) the no-broadcasting, (c) the impossibility of unconditionally secure bit commitment—they have shown that the algebra must be necessarily non-abelian.

In the recent article [9] one of the authors has proposed a set of postulates to be considered for a purely operational axiomatization of Quantum Theory, with the idea in the background that Quantum Theory is a set of coherence rules for managing probabilities in a game-theoretical *fair operational framework*. More precisely, Quantum Theory is a set of rules which allows the experimenter to make predictions on future *events* on the basis of suitable *tests* without interference from uncontrollable sources, with local control and with low experimental complexity. This is very close to the original spirit of Ludwig [6]. In addition to causality, the following postulates have been considered: PFAITH (existence of a pure preparationally faithful state), and FAITHE (existence of a faithful effect). These postulates have exhibited an unexpected theoretical power, excluding all known nonquantum probabilistic theories, such as PR-boxes [10], *rebits* [11], etc. The two postulate alone are however not sufficient to derive Quantum Theory, and other potential postulates of the same nature have been then considered, such as FAITHE: (existence of a faithful effect), SUPERFAITH (existence of a pure preparationally state which used in many copies also provides a $2n$-partite preparationally faithful states), and PURIFY-1 (purifiability of all states). More recently Chiribella et al. [12] presented a thorough axiomatic analysis, based on causality and postulates LDISCR (local discriminability) and PURIFY-2 (purifiability of all states, uniquely up to reversible channels on the purifying system). These postulates make the probabilistic framework much closer to Quantum Theory, with teleportation, error correction, dilation theorems, no cloning, and no bit commitment among its corollaries.

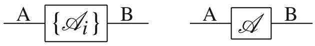
In the present paper we test the above postulates on basis of the existing probabilistic models that are not quantum. The first model, *the two-box world*, is an extension of the Popescu-Rohrlich model [10], which achieves the greatest violation of the CHSH inequality compatible with the no-signaling principle. The second model, *the two-clock world*, is actually a full class of models, all having a disk as convex set of states for the local system. These models allow purification of all its mixed states (PURIFY-1), but the purification is not unique up to reversible channels on the purifying system, as PURIFY-2 requires. One of the models of this class is indeed the *the two-rebit world*, namely qubits with real Hilbert space. This model violates LDISCR and then the local observability principle, namely the possibility of discriminating joint states by local measurements. The third model, *the spin-factor*, is a sort of n-dimensional generalization of the clock. Here we show that the only dimension $n = 3$ allows teleportation, and, indeed, in such case the theory is the *qubit*. Finally the last model is *the classical probabilistic theory* which are revisited in the introduced probabilistic theories framework. In pointing out which postulates are violated by each model we will also emphasize the deep relations that exist between the local and the non-local structures of the probabilistic theories.

The world of probabilistic theories is still largely unexplored, and we still have poor intuition, which, in addition, is also biased by our familiarity with Quantum Theory. Such lack of intuition is also a consequence of the absence of available alternative probabilistic models for testing new axiomatic frameworks, and this is the main motivation for the present paper.

2 Short review on probabilistic operational theories

2.1 The operational framework

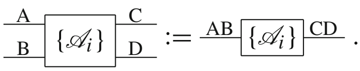
The primitive notion of our framework is the notion **test**. A test is made of the following ingredients: (a) a complete collection of **outcomes**, (b) input **system**, (c) output system. It is represented in form of a box, as follows



(1)

The left wire represents the input system, the right wire the output system, and $\{\mathscr{A}_i\}$ the collection of outcomes. Very often it is convenient to represent not the complete test, but just a single outcome $\mathscr{A}_i$, or, more generally a subset $\mathscr{A} \subset \{\mathscr{A}_i\}$ of the collection of possible outcomes, i.e. what is called **event**. The number of wires at the input and at the output can vary, and one can also have no wire at the input and/or at the output. We can regard the test in many different ways, depending on our needs and context. A test can be a man-made apparatus—such as a Stern-Gerlach setup or a beam splitter—or a nature-made "phenomenon"—such as a physical interaction between different particles in some space-time region. The set of events of a test is closed under union, intersection, and complementation, thus making a Boolean algebra. The **union** $\mathscr{A} \cup \mathscr{B}$ of two events $\mathscr{A}$ or $\mathscr{B}$ is the event in which either $\mathscr{A}$ or $\mathscr{B}$ occurred, but it is unknown which one. This operation is also called **coarse-graining**. Reversely, a **refinement** of an event $\mathscr{A}$ is a set of events $\{\mathscr{A}_i\}$ occurring in some test such that $\mathscr{A} = \cup_i \mathscr{A}_i$. Generally an event has different refinements, depending on the test, and is not refinable within some test. We will call an event that is unrefinable within any test **atomic event**.

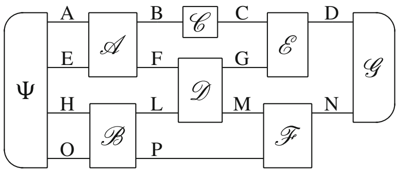
Composite system. Given two systems A and B we can consider them together as the **composite system** AB. A test $\{\mathscr{A}_i\}$ with input system AB and output system CD, represents one use of a physical device with two input and two output ports and diagrammatically it can be represented as the following box



(2)

In general the N-partite composite system $A_1 A_2 \ldots A_N$ will be represented with N wires.

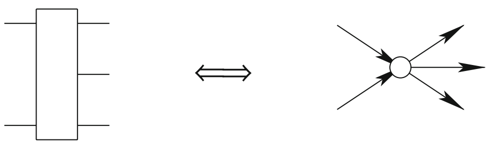
Connecting the test in a network. The natural place for a test/event will be inside a network of other tests/events, and to understand the origin of the box representation and the intimate meaning of the test/event you should imagine it actually connected to other tests/events in a circuit, e.g. as follows



The different letters A, B, C, ... labelling the wires precisely denote different "types of system", whose meaning comes from the following rules:

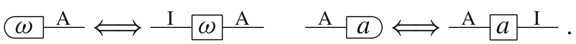
Connectivity rules: (1) we can connect only an input wire of a box with an output wire of another box, (2) we can connect only wires with the same label, (3) loops are forbidden.

Observation 2.1 *The fact that there are no closed loops gives to the circuit the structure of a* **directed acyclic graph** *(DAG). In the typical graph representation* **vertices** *correspond to operations, and* **edges** *to wires. The circuit and the graph representations are exactly equivalent, once one looks at a vertex as a "box" with inputs and outputs, as follows*



Ultimately the wires have only the function of ruling the way in which a box can be connected to other boxes. Thus *systems are just a representation of the causal connections between different events*. The fact that there are no closed loops corresponds to the requirement that the test/event is one-use only, whence each box in the circuit represents events that happen only once. Moreover, we must keep in mind that the probability of the event is independent on the test which it belongs, in the sense that if we have another test that contains the same event, this will have the same probability (keeping the rest of the network fixed). The fact that the probability depends only on the event and not on the test legitimates our use of networks made of single-event boxes, where on each box we don't need to specify the test. In the following, we will denote the set of events from system A to system B as $\mathfrak{T}(\mathrm{A}, \mathrm{B})$, and use the abbreviation $\mathfrak{T}(\mathrm{A}) := \mathfrak{T}(\mathrm{A}, \mathrm{A})$.

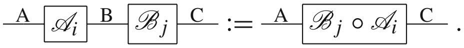
The trivial system. Among the different kinds of systems, we consider a special one called **trivial system**, denoted by I. In the circuit it will be represented by no wire, but instead we will draw the corresponding side of the operation box convexly rounded, namely as follows



(3)

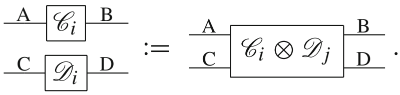
Building up the network formally. For convenience we can shortly point out the notions of parallel and sequential composition of tests. If we consider the test $\{\mathscr{A}_i\}_{i\in X}$

from A to B and the test $\{\mathscr{B}_j\}_{j\in Y}$ from B to C, we can take their sequential composition as the test from A to C with outcomes $(i, j) \in X \times Y$ and events $\{\mathscr{B}_j \circ \mathscr{A}_i\}_{(i,j)\in X\times Y}$. Diagrammatically the events $\mathscr{B}_j \circ \mathscr{A}_i$ are represented as follows



(4)

If $\{\mathscr{A}_i\}_{i\in X}$ is a test from A to B and $\{\mathscr{C}_j\}_{j\in Y}$ is a test from C to D then their parallel composition is the test from AC to BD with outcomes $(i, j) \in X \times Y$ and events $\{\mathscr{A}_i \otimes \mathscr{C}_j\}_{(i,j)\in X\times Y}$. Diagrammatically the events $\mathscr{A}_i \otimes \mathscr{C}_j$ and the events $\mathscr{B}_j \circ \mathscr{A}_i$ are represented as follows



(5)

One can build up the network using formal rules as in Chiribella et al. [12], making connection in parallel, in sequence, declaring commutativity of parallel composition, remembering the existence of a trivial system (which play the role of identity in the systems composition) etc. This construction is mathematically equivalent to the construction of a *symmetric strict monoidal category*, and poses a strong bridge with the research line of Coecke and Abramsky [13]. We also must keep in mind that there are no constraints for disconnected parts of the network, namely they can be arranged freely as long as they are disconnected (this for example would not be true for a quaternionic quantum network).

2.2 The operational probabilistic theory

If you now want to make predictions about the occurrence probability of events based on your current knowledge, then you need a "theory" that assign probabilities to different events:[1] *An* **operational theory** *is specified by a collection of systems, closed under parallel composition, and by a collection of tests, closed under parallel/sequential composition and under randomization. The operational theory is probabilistic if every test from the trivial system to the trivial system is associated to a probability distribution of outcomes.*

Therefore a probabilistic theory provides us with the joint probabilities for all possible events in each box for any closed network, namely which has no input and no output system. The probability itself will be conveniently represented by the corresponding network of events. One is seldom interested in full joint probabilities, but, more often, in the joint probability of events in some given tests in the network, irrespective of events in all other tests. This will correspond to marginalize over the other tests. We will see how the evaluation of probabilities will be greatly simplified by the causality assumption and by the use of conditional states.

[1] Probabilities in the network can be introduced in a easy intuitive way, or in a more axiomatic way as Chiribella et al. [12].

Slices, preparations and observations. Up to now we have not distinguished tests/events with or without input and output wires. In fact all the tests/events have been considered as the same elementary device of a circuit. On the other hand in an operational probabilistic theory (as defined in the previous paragraph) can be introduced the notions of preparations and observations as particular kinds of tests/events in which respectively the input and the output system is the trivial one. The reason for this nomenclature is clarified in the following.

Two wires in a circuit are input-output **contiguous** if they are the input and the output of the same box. By following input-output contiguous wires in a circuit while crossing boxes only in an input-output direction we draw an **input-output chain**. Two systems (wires) that are not in the same input-output path are called **independent**. A set of pairwise independent systems/wires is a **slice**, and the slice is called global if it partitions a closed bounded circuit into two parts as in Fig. 1 which, using our composition rules, is equivalent to the following

$$\left(\!\!\boxed{(\Psi_i, \mathscr{A}_j, \mathscr{B}_k, \mathscr{D}_m, \mathscr{F}_p)}\!\!\overset{\text{BGN}}{\text{———}}\!\!\boxed{(\mathscr{C}_l, \mathscr{E}_n, \mathrm{G}_q)}\!\!\right) \tag{6}$$

namely, it is equivalent to the connection of a **preparation test** with an **observation test**. Thus, a diagram of the form $(\!\boxed{\mathscr{A}_i}\overset{\text{A}}{\text{——}}\boxed{\mathscr{B}_j}\!)$ generally represents the event corresponding to an instance of a concluded experiment, which starts with a preparation and ends with an observation. The probability of such event will be denoted as $(\mathscr{B}_j \,|\, \mathscr{A}_i)$, using the "Dirac-like" notation, with *rounded ket* $|\mathscr{A}_i)$ and *rounded bra* $(\mathscr{B}_j|$ for the preparation and the observation tests, respectively. In the following we will use lowercase Greek letters for preparation-events and lowercase Latin letters for observation events. The following equivalent notations denote the probability of the sequence of events ρ, $\mathscr{A}$, a

$$(a|\,\mathscr{A}\,|\rho) = (\!\boxed{\rho}\!-\!\boxed{\mathscr{A}}\!-\!\boxed{a}\!). \tag{7}$$

It is also possible to denote the sequence as,

$$(\!\boxed{\rho}\!-\!\boxed{\mathscr{A}}\!-\!\boxed{a}\!) = (\!\boxed{\rho}\!-\!\boxed{a \circ \mathscr{A}}\!) \quad \text{with} \quad (a|\,\mathscr{A} = (a \circ \mathscr{A}|\,, \tag{8}$$

and the event $\mathscr{A}$ can be regarded as "transforming" the observation event a into the new observation-event $a \circ \mathscr{A}$. The same can be said for the preparation-event ρ which should have been regarded as the new transformed preparation-event $|\mathscr{A} \circ \rho)$.

2.3 States, effects, transformations

We will denote by $\mathfrak{S}(\mathrm{A})$ and $\mathfrak{E}(\mathrm{A})$ the set of all possible preparation and observation-events of a system A. In a probabilistic theory, a preparation-event ρ_i for system A is naturally identified with a function sending observation-events of A to probabilities, namely,

$$\rho_i : \mathfrak{E}(\mathrm{A}) \to [0, 1], \quad (a_j| \mapsto (a_j \,|\, \rho_i), \tag{9}$$

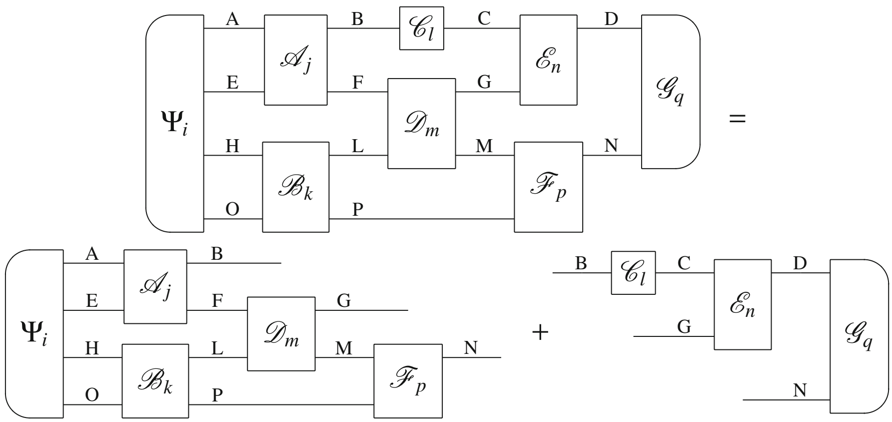


Fig. 1 Split of a closed circuit into a preparation and an observation test

and, analogously, observation-events are identified with functions from preparation-events to probabilities

$$a_j : \mathfrak{S}(\mathrm{A}) \to [0,1], \quad |\rho_i) \mapsto \left(a_j \,|\, \rho_i\right). \tag{10}$$

According to this picture a preparation-event is a probability rule for each possible observation-event on the system $\mathscr{A}$. If regarded as probability rules, two preparation-events corresponding to the same function are indistinguishable. The same can be said for observation-events regarded as probability rules over preparation-events. Thus given a system A we get

$$\begin{aligned} (a_i| = \left(a_j\right| &\Longleftrightarrow (a_i \,|\, \rho_k) = \left(a_j \,|\, \rho_k\right) \quad \forall\, |\rho_k) \in \mathfrak{S}(\mathrm{A}) \\ |\rho_i) = \left|\rho_j\right) &\Longleftrightarrow (a_k \,|\, \rho_i) = \left(a_k \,|\, \rho_j\right) \quad \forall\, (a_k| \in \mathfrak{E}(\mathrm{A}). \end{aligned} \tag{11}$$

We are thus lead to the following notions of states and effects in terms of equivalence classes of preparation/observation-events:

States and effects: *Equivalence classes of indistinguishable preparation-events for system* A *are called* **states** *of* A*, and their set is denoted as* $\mathfrak{S}(\mathrm{A})$. *Equivalence classes of indistinguishable observation-events for system* A *are called* **effects** *of* A*, and their set is denoted as* $\mathfrak{E}(\mathrm{A})$.

In the following we will make the identifications: (1) preparation-events $\equiv$ states; (2) observation-events $\equiv$ effects. Notice that according to our definition of states and effects as equivalence classes, states are separating for effects and viceversa effects are separating for states.[2]

[2] We say that a set of effects is **separating** for a set of states, if any two states of the set have at least a different probability for two effects of the other set. Similarly for a set of states.

Linear spaces of states and effects. Since states (effects) are functions from effects (states) to probabilities, one can take linear combinations of them. This defines the real vector spaces $\mathfrak{S}_{\mathbb{R}}(\mathrm{A})$ and $\mathfrak{E}_{\mathbb{R}}(\mathrm{A})$, one dual of the other (we will restrict our attention to finite dimensions). In this case, by duality one has $\dim(\mathfrak{S}_{\mathbb{R}}(\mathrm{A})) = \dim(\mathfrak{E}_{\mathbb{R}}(\mathrm{A}))$.

Convex cones of states and effects. Linear combinations with positive coefficients of states or of effects define the two convex cones $\mathfrak{S}_{+}(\mathrm{A})$ and $\mathfrak{E}_{+}(\mathrm{A})$, respectively, one dual cone of the other. The standard assumption in the literature is that, since the experimenter is free to randomize the choice of devices with arbitrary probabilities, the set of states $\mathfrak{S}(\mathrm{A})$ and the set of effects $\mathfrak{E}(\mathrm{A})$ are convex. In the following we will denote the extremal points of a convex set—say the $\mathfrak{S}(\mathrm{A})$—as $\mathrm{Extr}(\mathfrak{S}(\mathrm{A}))$. Naturally the points in $\mathrm{Extr}(\mathfrak{S}(\mathrm{A}))$ correspond to the pure states of the theory.

Linear extension of events. Linearity is naturally transferred to any kind of event through Eqs. (7) and (8), via linearity of probabilities, and, in addition, events become linear maps on states or effects, e.g. $\mathscr{A} \in \mathfrak{T}(\mathrm{A},\mathrm{B})$, $\mathscr{A} : |\rho)_{\mathrm{A}} \mapsto |\mathscr{A}\rho)_{\mathrm{B}}$. Every event $\mathscr{A} \in \mathfrak{T}(\mathrm{A},\mathrm{B})$ induces a map from $\mathfrak{S}(\mathrm{AC})$ to $\mathfrak{S}(\mathrm{BC})$ for every system C, uniquely defined by

$$\mathscr{A} : |\rho)_{\mathrm{AC}} \in \mathfrak{S}(\mathrm{AC}) \mapsto (\mathscr{A} \otimes \mathscr{I}_{\mathrm{C}})\,|\rho)_{\mathrm{AC}} \in \mathfrak{S}(\mathrm{BC}), \tag{12}$$

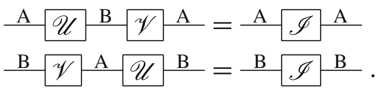
$\mathscr{I}_{\mathrm{C}}$ denoting the identity transformation on system C and $\otimes$ denoting the parallel composition of events as defined in Eq. (5). The map is linear from $\mathfrak{S}_{\mathbb{R}}(\mathrm{AC})$ to $\mathfrak{S}_{\mathbb{R}}(\mathrm{BC})$. From a probabilistic point of view, if for every possible system C two events $\mathscr{A}$ and $\mathscr{A}'$ induce the same maps, then they are indistinguishable. We are thus lead to the definition of **transformation**: *Equivalence classes of indistinguishable events from* A *to* B *are called* **transformations** *from* A *to* B. Henceforth, we will identify events with transformations. Accordingly, a test will be a collection of transformations.

In the following, if there is no ambiguity, we will drop the system index to the identity event. Notice that generally two transformations $\mathscr{A}, \mathscr{A}' \in \mathfrak{T}(\mathrm{A},\mathrm{B})$ can be different even if $\mathscr{A}\,|\rho)_{\mathrm{A}} = \mathscr{A}'\,|\rho)_{\mathrm{A}}$ for every $\rho \in \mathfrak{S}(\mathrm{A})$. Indeed one has $\mathscr{A} \neq \mathscr{A}'$ if there exists an ancillary system C and a joint state $|\rho)_{\mathrm{AC}}$ such that

$$(\mathscr{A} \otimes \mathscr{I})\,|\rho)_{\mathrm{AC}} \neq (\mathscr{A}' \otimes \mathscr{I})\,|\rho)_{\mathrm{AC}}. \tag{13}$$

We will come back on this point when discussing local discriminability.

Channels and automorphisms. A deterministic transformation $\mathscr{U} \in \mathfrak{T}(\mathrm{A},\mathrm{B})$ is called **channel**. A channel $\mathscr{U} \in \mathfrak{T}(\mathrm{A},\mathrm{B})$ is reversible if there is another channel $\mathscr{V} \in \mathfrak{T}(\mathrm{B},\mathrm{A})$ such that



(14)

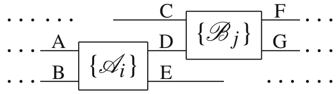
The set of **automorphisms** of a system A, the reversible channels from A to itself, form a group here denoted by $\mathbf{G}_{\mathrm{A}}$.

2.4 No signaling from the "future"

Although in the networks discussed until now we had sequences of tests, such sequences were not necessarily temporal, or *causal sequences*, namely the order of tests in a sequence was not necessarily following the causal or the time arrow.

We now introduce the causality condition, also called *no signalling from the future*, which allows us to interpret the sequential composition as a causal cascade.

Causality condition 1. (See also [9]) *We say that a theory is causal, if for any two tests* $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$ *and* $\{\mathscr{B}_j\}_{J\in\mathrm{Y}}$ *that are connected with at least an input of test* $\{\mathscr{B}_j\}_{J\in\mathrm{Y}}$ *connected to an output of* $\{\mathscr{A}_i\}_{J\in\mathrm{Y}}$ *as follows*



(15)

one has the asymmetry of the joint probability of events (given all other events in the network):

$$\sum_{j\in\mathrm{Y}} p(\mathscr{A}_i,\mathscr{B}_j) = p(\mathscr{A}_i), \quad \forall\mathscr{A}_i, \forall\{\mathscr{B}_j\}_{j\in\mathrm{Y}}, \tag{16}$$

$$\sum_{i\in\mathrm{X}} p(\mathscr{A}_i,\mathscr{B}_j) = p(\mathscr{B}_j,\{\mathscr{A}_i\}_{i\in\mathrm{X}}), \quad \forall\{\mathscr{A}_i\}_{i\in\mathrm{X}}, \forall\mathscr{B}_j. \tag{17}$$

In words, we say that the marginal over test $\{\mathscr{B}_j\}_{j\in\mathrm{Y}}$ *is independent on the choice of the same test—namely it would be the same if there were no test at the output of test* $\{\mathscr{A}_i\}_{j\in\mathrm{X}}$*—whereas the marginal over the test* $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$ *generally depends on the choice of the test.*

Causality and the arrow of time. The above asymmetry of marginalization of joint probabilities corresponds to say that: *test* $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$ *can influence test* $\{\mathscr{B}_j\}_{j\in\mathrm{Y}}$*, but not viceversa*. Or else: $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$ *is cause for* $\{\mathscr{B}_j\}_{j\in\mathrm{Y}}$ *and* $\{\mathscr{B}_j\}_{i\in\mathrm{Y}}$ *is effect for* $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$. Thus, the asymmetry is **causality**. If we now take the input-output direction as the past-future time relation, this corresponds to choose the **arrow of time**, namely it corresponds to say that causes precede effects. According to our choice of the time-arrow the input-output connection between tests is interpreted as a **time-cascade of tests**. Therefore, in synthesis, the asymmetry in the marginalization of probabilities corresponds to postulate that:

Postulate NSF: No signaling from the future. *The marginal probability of a time-cascade of tests does not depend on the "future" tests.*

On the contrary, the marginal probability of a time-cascade of tests generally depends on "past" tests, and we will see that this leads to the customary probability-conditioning from the past.

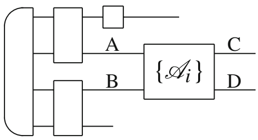
The causality condition greatly simplifies the evaluation of probabilities of events. In fact, since the probability of an event in a test is independent on the tests performed at the output, we can just substitute the network with another one in which all output systems of the test of interest are substituted by a deterministic test.

Formulation in terms of preparation tests. We have already introduced preparation tests, namely tests with no input, and denoted as $(\rho_i \overset{\mathrm{B}}{\vdash}$. Moreover, we have shown that every portion of network that has no input is equivalent to a preparation test, as e.g. in Fig. 1. Naturally this portion of circuit is the state of the system (in general composed) made of the free right wires of the circuit's portion under consideration. The causal condition can now be equivalently formulated as follows:

Causal Condition 2. [12] *A theory is causal if every preparation-event* $|\rho_j)_\mathrm{A}$ *has a probability* $p(\rho_j)$ *that is independent on the choice of test following the preparation test. Precisely, if* $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$ *is an arbitrary test from* A *to* B*, one has*

$$p(\rho_j) = \sum_{i\in\mathrm{X}} p(\mathscr{A}_i \rho_j). \tag{18}$$

The equivalence of the two formulations of the causal condition can be easily proved as follows. The implication Condition 1 $\Longrightarrow$ Condition 2 is immediate. Viceversa, consider any portion of the complete network which has no input, which contains test $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$, and which has noting attached at the output systems of test $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$, as follows



(19)

This is a preparation test. Then according to Condition 2 the joint probability of all events in the preparation test—i.e. our portion of network—is independent on the choice of tests connected at the output of the network. In particular, the probability of events $\mathscr{A}_i$ given all other events in the network will be independent on the choice of the test at the output of test $\{\mathscr{A}_i\}$.

We should emphasize that there exist indeed input-output relations that have no causal interpretation. Such non causal theories are studied in Chiribella et al. [14]. A concrete example of such theories is that considered in Refs. [15,16], where the states are quantum operations, and the transformations are "supermaps" transforming quantum operations into quantum operations. In this case, transforming a state means inserting the quantum operation in a larger circuit, and the sequence of two transformation is not a causal. The possibility of formulating more general probabilistic theories even in the absence of a pre-defined causal arrow may constitute a crucial ingredient for conceiving a quantum theory of gravity (see e.g. Hardy in Ref. [17]).

The causality principle naturally leads to the notion of conditioned tests, generalizing both notions of sequential composition and randomization of tests. For a precise definition see Ref. [12].

Causal theories have a simple characterization in terms of the following lemmas whose proof is in [12].

Lemma 1 *A theory is causal if and only if for every system* A *there is a unique deterministic effect* $(e|_{\mathrm{A}}$.

Lemma 2 *A theory where every state is proportional to a deterministic one is causal.*

No signaling without exchanging systems. The "no signalling from the future", i.e. the causality requirement, implies another "no signaling", namely the impossibility of signalling without exchanging systems:

Theorem 1 *(No signalling without exchange of physical systems) In a causal theory it is impossible to have signalling without exchanging systems.*

Proof See Ref. [12]. □

2.5 Alternative definition of *state* for causal theories

From Lemma 1 it is clear that in a causal theory the probability function over events p is uniquely defined. We can accordingly define the state also in the following way: *A* **state** ω *for a system* A *is a probability rule* $\omega(\mathscr{A})$ *for any event* $\mathscr{A} \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$ *occurring in any possible test with input system* A. *We call the state* **normalized** *if for every possible test* $\{\mathscr{A}_i\}_{i\in\mathrm{X}}$ *with input system* A, *the following condition holds*

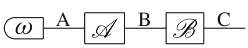
$$\sum_{j\in\mathrm{X}} \omega(\mathscr{A}_j) = 1. \tag{20}$$

It is easy to see that for causal theories the above definition is equivalent to the definition of state as equivalence class of preparation-event. In fact, the preparation-event is a positive functional over observation tests (see Eq. 9). On the other hand, due to causality, the probability of the event $\mathscr{A}$ for preparation $|\omega)_{\mathrm{A}}$ is independent on the choice of the following test, whence, in particular, is given by

$$\omega(\mathscr{A}) = {}_{\mathrm{B}}(e|\,\mathscr{A}\,|\omega)_{\mathrm{A}}, \tag{21}$$

whereas normalization follows easily. Viceversa, for a normalized state the probability rule $\omega(\mathscr{A})$ along with normalization (20) provides probabilities that satisfy Eq. (18).

Conditional state. Causality also allows us to define the notion of conditional state, namely the state corresponding to the conditional probability rule. The following cascade



(22)

leads to the notion of conditional probability that event $\mathscr{B}$ occurs knowing that event $\mathscr{A}$ has occurred $p(\mathscr{B}|\mathscr{A}) = \omega(\mathscr{B} \circ \mathscr{A})/\omega(\mathscr{A})$. This sets the new probability rule $\omega_{\mathscr{A}}(\mathscr{B}) := p(\mathscr{B}|\mathscr{A})$, corresponding to the notion of **conditional state**: *The conditional state $\omega_{\mathscr{A}}$, which gives the probability that an event occurs knowing that event $\mathscr{A}$ has occurred with the system prepared in the state ω, is given by*

$$\omega_{\mathscr{A}} \doteq \frac{\omega(\cdot \circ \mathscr{A})}{\omega(\mathscr{A})} \tag{23}$$

(the central dot "·" denoting the location of the variable). This is another way of regarding the event $\mathscr{A}$ as a transformation, namely as transforming with probability $\omega(\mathscr{A})$ the state ω to the (unnormalized) state $\mathscr{A}\omega$ given by

$$\mathscr{A}\omega := \omega(\cdot \circ \mathscr{A}). \tag{24}$$

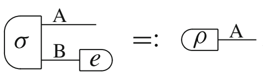
In such way causality leads to the identifications: (1) event $\equiv$ transformation and (2) evolution $\equiv$ state-conditioning. Notice that also a deterministic event produces a nontrivial conditioning of probabilities.

Marginal state. Regarding the state as a probability rule in causal theories naturally leads to the other relevant notion of **marginal state**, corresponding to the marginalization probability rule. The marginal state is just the probability rule for marginal probability, namely

$$\begin{aligned} p_{ij} &= (a_i|_{\mathrm{A}} \left(b_j\right|_{\mathrm{B}} |\sigma)_{\mathrm{AB}}, \\ p_i &= \textstyle\sum_j p_{ij} = \sum_j (a_i|_{\mathrm{A}} \left(b_j\right|_{\mathrm{B}} |\sigma)_{\mathrm{AB}} = (a_i|_{\mathrm{A}} (e|_{\mathrm{B}} |\sigma)_{\mathrm{AB}} =: (a_i|_{\mathrm{A}} |\rho)_{\mathrm{A}}. \end{aligned} \tag{25}$$

Here $|\rho)_{\mathrm{A}}$ is the marginal state of system A of the joint state $|\sigma)_{\mathrm{AB}}$. The definition of marginal state is therefore the following

Marginal state: *The marginal state of $|\sigma)_{\mathrm{AB}}$ on system* A *is the state $|\rho)_{\mathrm{A}} := (e|_{\mathrm{B}} |\sigma)_{\mathrm{AB}}$ and the marginalization of a state corresponds to the following diagram*



(26)

Abbreviated notation. In the following, when considering a transformation in $\mathscr{A} \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$ acting on a joint state $\omega \in \mathfrak{S}(\mathrm{AC})$, we will think the transformation acting on ω locally, namely we will use the following natural abbreviations

$$\mathscr{A} \in \mathfrak{T}(\mathrm{A}, \mathrm{B}),\ \omega \in \mathfrak{S}(\mathrm{AC}), \qquad \mathscr{A}\omega \equiv (\mathscr{A} \otimes \mathscr{I})\,|\omega)_{\mathrm{AC}}, \tag{27}$$

$$\omega(\mathscr{A}) \equiv (e|_{\mathrm{AC}}\,(\mathscr{A} \otimes \mathscr{I})\,|\omega)_{\mathrm{AC}}. \tag{28}$$

In regards of probabilities the abbreviation corresponds to take the marginal state.

Complete operational specification of a transformation. Operationally a transformation/event $\mathscr{A} \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$ needs to be completely specified by the way it affects all observed outcomes, namely all probabilities. This means that it is specified by all the joint probabilities in which it is involved. It follows that $\mathscr{A} \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$ is univocally given by the probability rule

$$\mathscr{A} \in \mathfrak{T}(\mathrm{A}, \mathrm{B}), \quad \mathscr{A}\omega = \omega(\cdot \circ \mathscr{A}), \quad \forall \omega \in \mathfrak{S}(\mathrm{AC}) \tag{29}$$

namely its local action on all joint states for any ancillary extensions. This is equivalent to specify both the conditional state $\omega_{\mathscr{A}}$ and the probability $\omega(\mathscr{A})$ for all possible states ω, due to the identity

$$\mathscr{A}\omega = \omega(\mathscr{A})\omega_{\mathscr{A}}. \tag{30}$$

In particular the **identity transformation** $\mathscr{I}$ is completely specified by the rule $\mathscr{I}\omega = \omega$ for all states ω.

Linear space of events. We have seen that states inherit a linear structure from being functionals over effects, and viceversa the effects inherit a linear structure from being functionals over states. We can also regard the linear combination of two states as reflecting the linear combination of their respective probability rule. On the other hand, since transformations/events are fully specified by their action on states, they are also completely specified by their action over their linear space, hence they inherit their linear structure as follows

$$(a\mathscr{A} + b\mathscr{B})\omega := a\mathscr{A}\omega + b\mathscr{B}\omega, \qquad \forall a, b \in \mathbb{R}, \ \forall \omega \in \mathfrak{S}(\mathrm{AC}) \tag{31}$$

namely, $\forall \mathscr{A}, \mathscr{B} \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$, the linear combination of events $a\mathscr{A} + b\mathscr{B}$ is completely specified by its action over a generic state $\omega \in \mathfrak{S}(\mathrm{AC})$, action that is given by the linear combination of the two states $\mathscr{A}\omega, \mathscr{B}\omega \in \mathfrak{S}(\mathrm{AC})$. According to Eq. (31) we can take linear combinations of transformations defining the real vector space $\mathfrak{T}_{\mathbb{R}}(\mathrm{A}, \mathrm{B})$, and taking linear combinations with positive coefficients we get the cone of transformations $\mathfrak{T}_{+}(\mathrm{A}, \mathrm{B})$. Naturally the set of extremal rays of the cone $\mathsf{Erays}(\mathfrak{T}_{+}(\mathrm{A}, \mathrm{B}))$ contain the atomic transformations which have been defined as the events that are unrefinable within any test. Notice that both compositions $\circ$ and $\otimes$ are distributive with respect to addition leading to an algebraic structure for the set of transformations (see Ref. [9]).

2.6 Alternative definition of *effect* for causal theories

An effect is the equivalence class of transformations occurring with the same probability.

Indeed, if the two transformations $\mathscr{A}_1, \mathscr{A}_2 \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$ are probabilistically equivalent, one has $(e|_{\mathrm{A}} \mathscr{A}_1 |\omega)_{\mathrm{A}} = (e|_{\mathrm{A}} \mathscr{A}_2 |\omega)_{\mathrm{A}}, \quad \forall \omega \in \mathfrak{S}(\mathrm{A})$, and due to the fact that states are separating for effects, this is equivalent to the identity of effects $(e|_{\mathrm{A}} \mathscr{A}_1 =$

$(e|_{\mathrm{A}} \mathscr{A}_2 := (a|$, and we will say that the two transformations belong to the same effect $a \in \mathfrak{E}(\mathrm{A})$.

Depending on the context, in the following we will also use the equivalent notations for states, effects, and transformations

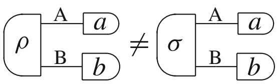
$$b \circ \mathscr{A} = (b| \mathscr{A}, \quad \mathscr{A}\omega = \mathscr{A}\,|\omega), \quad (b|\,\mathscr{A}\,|\omega) = \omega(b \circ \mathscr{A}). \tag{32}$$

Observables. One of the consequences of Lemma 1 is that the set of effects $\{l_i\}$ corresponding to all possible events of a test satisfies the normalization identity $\sum_i l_i = e$, e denoting the deterministic effect. A set of effects $\{l_i\}$ summing to the deterministic one is called **observable**. We will also call an observable **informationally complete** for A when each $a \in \mathfrak{E}(\mathrm{A})$ can be written as a real linear combinations of l_i, namely when the set of effects is separating for the set of states $\mathfrak{S}(\mathrm{A})$ (see footnote 2). If the effects l_i are all linearly independent then the informationally complete observable is said **minimal**. Beside the notion of informationally complete observable there exists an analogue notion of separating set of states for the set of effects.

2.7 Local discriminability

A standard assumption in the literature on probabilistic theories is *local discriminability* (LDISCR).

Local discriminability: *A theory satisfies local discriminability if for every couple of different states* $\rho, \sigma \in \mathfrak{S}(\mathrm{AB})$ *there are two local effects* $a \in \mathfrak{E}(\mathrm{A})$ *and* $b \in \mathfrak{E}(\mathrm{B})$ *such that*



(33)

Another way of stating local discriminability is to say that the set of factorized effects is separating for the joint states.

Local discriminability represents a dramatic experimental advantage. Without local discriminability, one generally would need to built up a N-system test in order to discriminate an N-partite joint state, instead of using just N of the same single system tests that allow us to discriminate states of single system. Local discriminability implies local observability, namely the possibility of recovering the full joint state from just local observations. Stated in other words, local observability means that one can build up an informationally complete observation test made only of local tests, i.e. one can perform a complete tomography of a multipartite state using only local tests. This is given by the following lemma [12]:

Lemma 3 *Let* $\{\rho_i\}$ *and* $\{\tilde{\rho}_j\}$ *be two bases for the vector spaces* $\mathfrak{S}_{\mathbb{R}}(\mathrm{A})$ *and* $\mathfrak{S}_{\mathbb{R}}(\mathrm{B})$*, respectively, and let* $\{a_i\}$ *and* $\{b_j\}$ *be two bases for the vector spaces* $\mathfrak{E}_{\mathbb{R}}(\mathrm{A})$ *and* $\mathfrak{E}_{\mathbb{R}}(\mathrm{B})$*, respectively. Then every state* $\sigma \in \mathfrak{S}(\mathrm{AB})$ *and every effect* $E \in \mathfrak{E}(\mathrm{AB})$ *can* *be written as follows*

$$\begin{aligned} |\sigma)_{\mathrm{AB}} &= \textstyle\sum_{i,j} \; A_{ij} \, |\rho_i)_{\mathrm{A}} \, \big|\tilde{\rho}_j\big)_{\mathrm{B}} \\ (E|_{\mathrm{AB}} &= \textstyle\sum_{i,j} \; B_{ij} \, (a_i|_{\mathrm{A}} \, \big(b_j\big|_{\mathrm{B}} \end{aligned} \tag{34}$$

for some suitable real matrix A_{ij} (B_{ij}).

Another consequence of local discriminability is that transformations in $\mathfrak{T}(\mathrm{A}, \mathrm{B})$ are completely specified by their action only on local states $\mathfrak{S}(\mathrm{A})$, without the need of considering ancillary extension. This is assessed by the following lemma [12]:

Lemma 4 *If two transformations $\mathscr{A}, \mathscr{B} \in \mathfrak{T}(\mathrm{A}, \mathrm{B})$ are different and local discriminability holds, then there exists a state $\rho \in \mathfrak{S}(\mathrm{A})$ such that*

$$(\rho|\!\!-^{\mathrm{A}}\!\!\boxed{\mathscr{A}}\!\!- \;\neq\; (\rho|\!\!-^{\mathrm{A}}\!\!\boxed{\mathscr{B}}\!\!- . \tag{35}$$

3 The Postulates PFAITH, FAITHE, and PURIFY-1

3.1 Postulate PFAITH

Postulate PFAITH plays a major role in the operational probabilistic theories of Refs. [9] and [12]. The Postulate concerns the possibility of calibrating any test and of preparing any joint bipartite state only by means of local transformations. Before introducing the Postulate we need to define what is a faithful state.

Consider a bipartite system AB and a bipartite state $\Phi \in \mathfrak{S}(\mathrm{AB})$. The state Φ induces the following cone-homomorphism[3]

$$\mathfrak{T}_+(\mathrm{A}) \ni \mathscr{A} \mapsto (\mathscr{A} \otimes \mathscr{I})\Phi \in \mathfrak{S}_+(\mathrm{AB}). \tag{36}$$

- If the cone-homomorphism in Eq. (36) is a cone-monomorphism, namely the output $(\mathscr{A} \otimes \mathscr{I})\Phi$ is in one to one correspondence with the local transformations $\mathscr{A}$, then Φ is **dinamically faithful** with respect to A. The output keeps the information about the input transformation and this allows to calibrate any test by means of local transformations.
- If the cone-homomorphism in Eq. (36) is a cone-epimorphism, namely every bipartite state Ψ can be achieved as $\Psi = (\mathscr{A}_\Psi \otimes \mathscr{I})\Phi$ for some local transformation $\mathscr{A}_\Psi$, then Φ is **preparationally faithful** with respect to A. Any joint state can be prepared by means of local transformations.

Observation 3.1 *For Φ both preparationally and dynamically faithful, one can operationally define the* **transposed transformation** *$\mathscr{A}' \in \mathfrak{T}_{\mathbb{R}}(\mathrm{A})$ of a transformation $\mathscr{A} \in \mathfrak{T}_{\mathbb{R}}(\mathrm{A})$ through the identity*

$$(\mathscr{A}' \otimes \mathscr{I})\Phi = (\mathscr{I} \otimes \mathscr{A})\Phi. \tag{37}$$

[3] A cone-homomorphism between the cones K_1 and K_2 is simply a linear map between $\mathsf{Span}_{\mathbb{R}}(K_1)$ and $\mathsf{Span}_{\mathbb{R}}(K_2)$ which sends elements of K_1 to elements of K_2, but not necessarily viceversa.

All the properties of transposition are verified.

Postulate PFAITH: Existence of a symmetric preparationally faithful pure state. *For any couple of identical systems, there exists a symmetric (under permutation of the two systems) bipartite state which is both pure and preparationally faithful.*

Postulate PFAITH leads to many relevant features of the probabilistic theory. Here we briefly report those that are useful in the construction of our concrete probabilistic models. For the proof see Ref. [9] where many other consequences are investigated. In the following, when considering two identical systems A = B if there is no ambiguity we will just write AA instead of AB to denote the bipartite system. Consider a probabilistic theory for two identical systems A = B that satisfies Postulate PFAITH and let Φ be a pure symmetric and preparationally faithful bipartite state of the theory; then the following properties hold [9]:

(1) Φ is both preparationally and dinamically faithful with respect to both systems.
(2) One has the cone-isomorphism[4] $\mathfrak{T}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{AA})$ induced by Φ via the map $\mathscr{A} \in \mathfrak{T}_+(\mathrm{A}) \leftrightarrow (\mathscr{A} \otimes \mathscr{I})\Phi \in \mathfrak{S}_+(\mathrm{AA})$. Moreover, a local transformation on Φ produces an output pure (unnormalized) bipartite state if and only if the transformation is atomic, namely $\Psi = (\mathscr{A}_\Psi \otimes \mathscr{I})\Phi$ is pure if and only if $\mathscr{A}_\Psi$ is atomic.
(3) The theory is weakly self-dual, namely one has the cone-isomorphism $\mathfrak{E}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{A})$ induced by the map $\Phi(a, \cdot) = \omega_a \; \forall a \in \mathfrak{E}_+(\mathrm{A})$.
(4) The identical transformation $\mathscr{I}$ is atomic.
(5) The transposed of an automorphism is still an automorphism.
(6) The maximally chaotic state $\chi := \Phi(e, \cdot)$ is invariant under the transpose of a channel (deterministic transformation) whence, in particular, under the group of automorphisms.

Observation 3.2 *A stronger version of PFAITH, satisfied by Quantum Theory, requires the existence of a symmetric preparationally* **superfaithful** *state Φ, such that also $\Phi \otimes \Phi$ is preparationally faithful, whence $\Phi^{\otimes 2n}$ is preparationally faithful with respect to A^n, $\forall n > 1$.*

3.2 Postulate FAITHE and teleportation

In Ref. [9] other Postulates are introduced which make the probabilistic theories closer to Quantum Theory. In this paper that Postulates will be tested on concrete probabilistic models.

Since a preparationally faithful state is also dynamically faithful, it is indeed an isomorphism, it is invertible. On the other hand, in general its inverse is not a bipartite effect:

[4] Two cones K_1 and K_2 are isomorphic if and only if there exists a linear bijective map between the linear spans $\mathsf{Span}_{\mathbb{R}}(K_1)$ and $\mathsf{Span}_{\mathbb{R}}(K_2)$ that is cone preserving in both directions, namely it and its inverse map must send $\mathsf{Erays}(K_1)$ to $\mathsf{Erays}(K_2)$ and positive linear combinations to positive linear combinations.

Postulate FAITHE: Existence of a faithful effect. *There exists a bipartite effect $F \in \mathfrak{E}(\mathrm{AA})$ achieving probabilistically the inverse of the cone-isomorphism $\mathfrak{E}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{A})$ given by $a \to \omega_a := \Phi(a, \cdot)$, namely*

$$(F|_{23}\,|\omega_a)_2 = (F|_{23}\,(a|_1\,|\Phi)_{12} = \alpha\,(a|_3\,, \quad 0 \leqslant \alpha \leqslant 1. \tag{38}$$

Equation (38) is equivalent to $(F|_{23}\,|\Phi)_{12} = \alpha \mathscr{S}_{13}$, $\mathscr{S}_{ij}$ denoting the transformation which swaps the ith system with the jth system. The main consequence of FAITHE is the possibility of achieving **probabilistic teleportation** of states between equal systems using the effect F and the state Φ as follows

$$\begin{aligned}(F|_{23}\,|\omega)_2\,|\Phi)_{34} &= (F|_{23}\,(a_\omega|_1\,|\Phi)_{12}\,|\Phi)_{34} = (a_\omega|_1\,\underbrace{(F|_{23}\,|\Phi)_{12}}_{\mathscr{S}_{13}}\,|\Phi)_{34} \\ &= \alpha\,(a_\omega|_1\,|\Phi)_{14} = \alpha\,|\omega)_4. \end{aligned} \tag{39}$$

According to the last equation Postulate FAITHE is equivalent to the relation

$$(F|_{23}\,|\Phi)_{12}\,|\Phi)_{34} = \alpha\,|\Phi)_{14}, \tag{40}$$

where α is the probability of achieving teleportation. It depends only on the faithful effect F since it is $\alpha = (e|_{14}\,(F|_{23}\,|\Phi)_{12}\,|\Phi)_{34}$. Moreover, the maximum value of α is achieved maximizing over all bipartite effects and it depends on the particular probabilistic theory.

Here we give a criterion to exclude the possibility of achieving teleportation from a preparationally faithful state in a probabilistic theory.

Proposition 3.1 *If there exists a preparationally faithful state violating Postulate FAITHE then all the preparationally faithful states violate it.*

Proof Let $\Phi \in \mathfrak{S}(\mathrm{AA})$ be the preparationally faithful state violating FAITHE. And let $F = \alpha\Phi^{-1}$ be the bipartite functional satisfying Eq. (38). Then there exists a state $\Psi \in \mathfrak{S}(\mathrm{AA})$ such that

$$\left(\Phi^{-1}\,|\,\Psi\right) < 0, \tag{41}$$

namely $F = \alpha\Phi^{-1}$ is not a true effect for each $\alpha > 0$. Now let Φ' be another preparationally faithful state. From the faithfulness of Φ, there exists a transformation $\mathscr{A} \in \mathfrak{T}(\mathrm{A})$ such that

$$(\mathscr{A} \otimes \mathscr{I})\Phi = \Phi'. \tag{42}$$

Φ' is preparationally and dinamically faithful, therefore $\mathscr{A}$ is invertible and $\mathscr{A}^{-1}$ is a transformation. Consider then the quantity

$$\left(\Phi^{-1}\,|\,\Psi\right) = \left(\Phi^{-1}\right|(\mathscr{A} \otimes \mathscr{I})(\mathscr{A}^{-1} \otimes \mathscr{I})\,|\Psi) = \left(\Phi'^{-1}\,|\,\Psi'\right) < 0. \tag{43}$$

So also Φ'^{-1} is not a bipartite effect because we have found a state $\Psi' = (\mathscr{A}^{-1} \otimes \mathscr{I})\Psi \in \mathfrak{S}(\mathrm{AA})$ such that $\left(\Phi^{-1\prime} \,|\, \Psi'\right) < 0$. □

As immediate consequence of this Proposition we get

Corollary 3.1 *If a probabilistic theory does not satisfy Postulate FAITHE then there is no preparationally faithful state achieving teleportation.*

The following Proposition will be useful in the construction of probabilistic models because it shows that a model which violates Postulate FAITHE cannot admit the existence of a super-faithful state.

Proposition 3.2 *If a probabilistic theory admits a super-faithful state Φ, then Postulate FAITHE is automatically satisfied and teleportation is achievable.*

Proof In fact considering the symmetric faithful quadripartite state $\Phi_{quad} = \Phi \otimes \Phi$, according to the isomorphism $\mathfrak{E}_{+}(\mathrm{AA}) \simeq \mathfrak{S}_{+}(\mathrm{AA})$, we can find a bipartite effect $F_{\Phi} \in \mathfrak{E}(\mathrm{AA})$ such that

$$(F_{\Phi}|_{23} \,|\Phi)_{12} \,|\Phi)_{34} = \alpha \,|\Phi)_{14}, \tag{44}$$

as required by FAITHE (see Eq. 40). Naturally teleportation follows as a consequence of Postulate FAITHE. □

3.3 Purifiability of a probabilistic theory

We know that Quantum Theory allows purification. A "minimal" version of purifiability for probabilistic theories is introduced through the following Postulate:

Postulate PURIFY-1: Purifiability of all states. *For every state $\omega \in \mathfrak{S}(\mathrm{A})$ there exists a system* B *and a pure state $\Omega \in \mathfrak{S}(\mathrm{AB})$ that purifies it, namely a state $\Omega \in \mathrm{Extr}(\mathfrak{S}(\mathrm{AB}))$ having ω as marginal state of system* A

$$(e|_{\mathrm{B}} \,|\Omega)_{AB} = |\omega)_{A}. \tag{45}$$

In Ref. [12] the numerous consequences of Postulate PURIFY-1 are analysed. In particular the following Lemma is proved.

Lemma 5 *If Postulate* PURIFY-1 holds then the identical transformation is atomic and the preparationally faithful state Φ is pure.

As already mentioned Postulate PURIFY-1 introduces a minimal notion of purifiability. Quantum Theory satisfies a more restrictive condition. Therefore in the same Ref. [12] a stronger version of Postulate PURIFY-1 is introduced:

Postulate PURIFY-2: Unique purification up to reversible channels on the purifying system. *Every state has a purification. If $\Omega_1 \in \mathfrak{S}(\mathrm{AB})$ and $\Omega_2 \in \mathfrak{S}(\mathrm{AB})$*

are both purifications of $\omega \in \mathfrak{S}(A)$*, then they are connected by a reversible local transformation* $\mathscr{U} \in \mathfrak{T}(\mathrm{B})$*, namely*

$$(e|_{\mathrm{B}} |\Omega_1)_{\mathrm{AB}} = (e|_{\mathrm{B}} |\Omega_2)_{\mathrm{AB}} = |\omega)_{\mathrm{A}} \Rightarrow |\Omega_2)_{\mathrm{AB}} = (\mathscr{I}_{\mathrm{A}} \otimes \mathscr{U}_{\mathrm{B}}) |\Omega_1)_{\mathrm{AB}}. \tag{46}$$

Postulate PURIFY-2 requires the uniqueness of the purification up to reversible channels on the purifying system at all the multipartite levels. Moreover it entails *entanglement swapping*, whence probabilistic teleportation:

Proposition 3.3 *If a probabilistic theory satisfies PURIFY-2, then each symmetric pure bipartite preparationally faithful state* $\Phi \in \mathfrak{S}(\mathrm{AA})$ *allows entanglement swapping. Thus the theory satisfies FAITHE and probabilistic teleportation is achievable.*

For the proof see Ref. [12]. Given a faithful state $\Phi \in \mathfrak{S}(\mathrm{AA})$ we say that the **entanglement swapping** is possible if there exists a constant $\alpha > 0$ and a bipartite effect $F \in \mathfrak{E}(\mathrm{AA})$ such that

$$(F|_{23} |\Phi)_{12} |\Phi)_{34} = \alpha |\Phi)_{14}. \tag{47}$$

Therefore, according to Eq. (40) FAITHE is satisfied and teleportation is achievable.

4 Bloch representation for transformations of a probabilistic theory

Based on the linear structure established for states, effects, and transformations, we can now introduce an affine-space representation based on the existence of a minimal informationally complete observable and of a separating set of states. Such representation generalizes the popular Bloch representation used in Quantum Theory.

In terms of a minimal informationally complete observable, $\{l_i\}$, $i = 1, \ldots, n$, and of a minimal separating set of states $\{\lambda_j\}$, $j = 1, \ldots, n$, one can expand (in a unique way) any effect $a \in \mathfrak{E}(\mathrm{A})$ and state $\omega \in \mathfrak{S}(\mathrm{A})$ as follows

$$a = \sum_{j=1}^{n} \lambda_j(a) l_j, \qquad \omega = \sum_{j=1}^{n} l_j(\omega) \lambda_j. \tag{48}$$

Instead of using a minimal informationally complete observable and a minimal set of separating states it is convenient to adopt canonical biorthogonal basis $\boldsymbol{l} = \{l_i\}$ and $\boldsymbol{\lambda} = \{\lambda_j\}$ for $\mathfrak{E}_{\mathbb{R}}(\mathrm{A})$ and $\mathfrak{S}_{\mathbb{R}}(\mathrm{A})$ embedded into $\mathbb{R}^n$ as Euclidean space, and identify an element in $\{l_i\}$ with the deterministic effect e—say $l_n = e$. Correspondingly λ_n in $\{\lambda_j\}$ is the functional χ giving the deterministic component of the effect. Using a Minkowskian notation

$$\boldsymbol{l} \doteq (\hat{\boldsymbol{l}}, e),\ \boldsymbol{\lambda} \doteq (\hat{\boldsymbol{\lambda}}, \chi) \qquad \text{with} \qquad \lambda \cdot l \doteq \sum_j \lambda_j l_j = \hat{\boldsymbol{\lambda}} \cdot \hat{\boldsymbol{l}} + \chi e, \tag{49}$$

we write

$$\begin{aligned}(a,\omega) &= \omega(a) = a(\omega)\\ &= \boldsymbol{l}(\omega)\cdot\boldsymbol{\lambda}(a) := \sum_{i=1}^{n} l_i(\omega)\lambda_i(a) \equiv \hat{\boldsymbol{\lambda}}(a)\cdot\hat{\boldsymbol{l}}(\omega)+\chi(a)e(\omega).\end{aligned} \tag{50}$$

The vectors $\boldsymbol{l}(\omega)$ and $\boldsymbol{\lambda}(a)$ give a complete description of the (unnormalized) state ω and (unbounded) effect a, thanks to identity (50). For a normalized state ω, $\boldsymbol{l}(\omega)$ is the **Bloch vector** representing the state ω.[5] For biorthogonal basis or, in general, for minimal informationally complete observables and separating set of states, the representation is *faithful* (i.e. one-to-one).

Consider a transformation $\mathscr{A} \in \mathfrak{T}(\mathrm{A})$. We now recover the linear transformation describing conditioning. The conditioning is given by $(b|\mathscr{A}|\omega) = \mathscr{A}\omega(b) = \omega(b\circ\mathscr{A}) = b(\mathscr{A}\omega)$. From the linearity of transformations one can introduce a matrix $\boldsymbol{A}\equiv\{A_{ij}\}$ such that $l_i\circ\mathscr{A} = \sum_j \mathscr{A}_{ij}l_j$, and then

$$l_i(\mathscr{A}\omega) = \omega(l_i\circ\mathscr{A}) = \sum_{j=1}^{n-1} A_{ij}l_j(\omega) + A_{in}e(\omega). \tag{51}$$

In particular, if $e\circ\mathscr{A} = a$, namely $\mathscr{A}$ is in the equivalence class a, one has

$$e(\mathscr{A}\omega) = \omega(e\circ\mathscr{A}) \equiv \omega(a) = \sum_{j=1}^{n} A_{nj}l_j(\omega) \equiv \hat{\boldsymbol{\lambda}}(a)\cdot\hat{\boldsymbol{l}}(\omega)+\chi(a)e(\omega), \tag{52}$$

from which we derive the identities $\lambda_j(a)\equiv A_{nj}$ and $\chi(a) = A_{nn}$.

The real matrices $\boldsymbol{A}$ are a *representation* of the real algebra of generalized transformations $\mathcal{A}$. The last row of the matrix is a representation of the effect a (see Fig. 2). In vector notation, for a normalized input state one has

$$\begin{aligned}\boldsymbol{l}(\mathscr{A}\omega) &= \hat{\boldsymbol{A}}\boldsymbol{l}(\omega)+\hat{\boldsymbol{k}}(\mathscr{A}), \qquad \hat{\boldsymbol{k}}(\mathscr{A}) \doteq \hat{\boldsymbol{l}}(\mathscr{A}\chi),\\ e(\mathscr{A}\omega) &= \hat{\boldsymbol{\lambda}}(a)\cdot\hat{\boldsymbol{l}}(\omega)+\chi(a),\\ \mathscr{A}\omega(b) &= \hat{\boldsymbol{\lambda}}(b)\cdot\hat{\boldsymbol{l}}(\mathscr{A}\omega)+\chi(b)e(\mathscr{A}\omega).\end{aligned} \tag{53}$$

The matrix representation of the transformation is given in Fig. 2.

Therefore, summarizing, we have found the following representation for the conditional state $\omega_{\mathscr{A}}$ after the action of the transformation $\mathscr{A}$ regarded as an affine map over $\mathfrak{S}(\mathrm{A})$

$$\omega\in\mathfrak{S}(\mathrm{A}),\quad \hat{\boldsymbol{l}}(\omega)\longrightarrow\hat{\boldsymbol{l}}(\omega_{\mathscr{A}}) = \frac{\hat{\boldsymbol{A}}\hat{\boldsymbol{l}}(\omega)+\hat{\boldsymbol{k}}(\mathscr{A})}{\hat{\boldsymbol{\lambda}}(a)\cdot\hat{\boldsymbol{l}}(\omega)+\chi(a)}, \tag{54}$$

[5] More precisely the last component of $\boldsymbol{l}(\omega)$ is $e(\omega) = 1$ for each normalized ω, and the Bloch vector is $\hat{\boldsymbol{l}}(\omega)$.

$$\boldsymbol{A} = \begin{pmatrix} \hat{\boldsymbol{A}} & \hat{\boldsymbol{k}}(\mathscr{A}) \\ \hat{\boldsymbol{\lambda}}(a)^t & \chi(a) \end{pmatrix}, \qquad \begin{aligned} \hat{\boldsymbol{l}}(\mathscr{A}\omega) &= \boldsymbol{A}\hat{\boldsymbol{l}}(\omega) + \hat{\boldsymbol{k}}(\mathscr{A}), \\ \omega(\mathscr{A}) &= \hat{\boldsymbol{\lambda}}(a) \cdot \hat{\boldsymbol{l}}(\omega) + \chi(a), \\ \hat{\boldsymbol{l}}(\omega_{\mathscr{A}}) &= \frac{\hat{\boldsymbol{A}}(\mathscr{A})\hat{\boldsymbol{l}}(\omega) + \hat{\boldsymbol{k}}(\mathscr{A})}{\hat{\boldsymbol{\lambda}}(a) \cdot \hat{\boldsymbol{l}}(\omega) + \chi(a)}. \end{aligned}$$

Fig. 2 Matrix representation of the real algebra of transformations $\mathcal{A}$. The last row represents the effect a of the transformation $\mathscr{A}$. It gives the transformation of the n-th component of the Bloch vector $e(\mathscr{A}\omega) \equiv \omega(\mathscr{A}) = \hat{\boldsymbol{\lambda}}(a) \cdot \hat{\boldsymbol{l}}(\omega) + \chi(a)$, namely the probability that $\mathscr{A}$ occurs. The other rows represent the affine transformation of the Bloch vector $\hat{\boldsymbol{l}}(\omega)$ corresponding to the action of $\mathscr{A}$, the last column giving the translation $\hat{\boldsymbol{k}}(\mathscr{A})$, and the remaining square matrix $\hat{\boldsymbol{A}}$ the linear part of the affine map. The Bloch vector of the state ω is transformed as $\hat{\boldsymbol{l}}(\mathscr{A}\omega) = \hat{\boldsymbol{A}}\hat{\boldsymbol{l}}(\omega) + \hat{\boldsymbol{k}}(\mathscr{A})$, and the conditioning over the convex set of states is the fractional affine transformation in figure

with the transformation occurring with probability $\omega(\mathscr{A})$ given by $\omega(\mathscr{A}) = \hat{\boldsymbol{\lambda}}(a) \cdot \hat{\boldsymbol{l}}(\omega) + \chi(a)$. One has

$$\omega \in \mathfrak{S}(\mathrm{A}), \quad \boldsymbol{l}(\omega) \longrightarrow \boldsymbol{l}(\omega_{\mathscr{A}}) = \left[\frac{\hat{\boldsymbol{A}}\hat{\boldsymbol{l}}(\omega) + \hat{\boldsymbol{k}}(\mathscr{A})}{\hat{\boldsymbol{\lambda}}(a) \cdot \hat{\boldsymbol{l}}(\omega) + \chi(a)}, 1 \right]. \tag{55}$$

A pictorially view of how the affine map $\mathscr{A}$ acts over $\mathfrak{S}(\mathrm{A})$ is given by the *linear-fractional map* and the *perspective map* (see [18]).

The following Propositions will be useful in constructing concrete probabilistic models.

Proposition 4.1 *A contraction in the convex set* $\mathfrak{T}(\mathrm{A})$ *is represented in Bloch form by a matrix having as last row an element of* $\mathfrak{E}(\mathrm{A})$. *Otherwise it cannot be a contraction.*

Proof By definition of Bloch representation. □

In the following we will denote by $\mathrm{Extr}(a)$ the set of all extremal transformations having effect a.

Proposition 4.2 *If* $a \in \mathrm{Extr}(\mathfrak{E}(\mathrm{A}))$ *and* $\mathscr{A} \in \mathrm{Extr}(a)$ *then* $\mathscr{A} \in \mathrm{Extr}(\mathfrak{T}(\mathrm{A}))$.

Proof If $\mathscr{A} \in a$ then its Bloch matrix has $\boldsymbol{\lambda}(a)$ as last row. According to Proposition 4.1 every set of contractions combining convexly to give $\mathscr{A}$ must combine to $\boldsymbol{\lambda}(a)$ in the last row of the Bloch representation. Since $a \in \mathrm{Extr}(\mathfrak{E}(\mathrm{A}))$ all the elements in the convex combination must be in the same equivalence class a in contradiction with the hypothesis $\mathscr{A} \in \mathrm{Extr}(a)$. □

Observation 4.1 *One could think that all extremal transformations are in the equivalence class of an extremal effect and that* $\mathrm{Extr}(\mathfrak{T}(\mathrm{A})) = \{\mathscr{A} \in \mathrm{Extr}(a), \forall a \in \mathrm{Extr}(\mathfrak{E}(\mathrm{A}))\}$. *In general this is false and Quantum Theory is a good counterexample. On the contrary, we will show how the extended Popescu-Rohrlich model satisfies this property.*

Definition We define the **generator set of** $\mathfrak{E}(\mathrm{A})$—denoted by $\mathsf{gen}(\mathfrak{E}(\mathrm{A}))$—as the set of effects whose orbit under the group of automorphisms for the system A is $\mathfrak{E}(\mathrm{A})$, namely the set such that $\mathsf{gen}(\mathfrak{E}(\mathrm{A})) \circ \mathbf{G}_{\mathrm{A}} = \mathfrak{E}(\mathrm{A})$.

Proposition 4.3 *The following equality holds*

$$\mathrm{Extr}(a \circ \mathbf{G}_{\mathrm{A}}) = \mathrm{Extr}(a) \circ \mathbf{G}_{\mathrm{A}} \quad \forall a \in \mathfrak{E}(\mathrm{A}) \tag{56}$$

Proof It is sufficient to show that $\forall a \in \mathsf{gen}(\mathfrak{E}(\mathrm{A}))$ and $\forall \mathscr{U} \in \mathbf{G}_{\mathrm{A}}$ one has

$$\mathrm{Extr}(a \circ \mathscr{U}) = \mathrm{Extr}(a) \circ \mathscr{U}. \tag{57}$$

Considering a map $\mathscr{B}$ in $\mathrm{Extr}(b)$ it is easy to show that $\mathscr{B} \circ \mathscr{U}$ is a map in $\mathrm{Extr}(b \circ \mathscr{U})$, in fact the Bloch representative of $\mathscr{B} \circ \mathscr{U}$ has $\boldsymbol{\lambda}(b \circ \mathscr{U})$ as last row and it is extremal because $\mathscr{B} = \mathscr{B} \circ \mathscr{U} \circ \mathscr{U}^{-1}$ is extremal. Then for each $\mathscr{A} \in \mathrm{Extr}(a \circ \mathscr{U})$ in Eq. (57) we can take $\mathscr{A} \circ \mathscr{U}^{-1} \in \mathrm{Extr}(a)$ satisfying the equality, and viceversa for each $\mathscr{A} \in \mathrm{Extr}(a)$ we can take $\mathscr{A} \circ \mathscr{U} \in \mathrm{Extr}(a \circ \mathscr{U})$. □

In the next sections we will construct some concrete probabilistic models, always considering composed identical systems, whence we will often omit the system specification in the expressions for the single system convex sets (e.g. $\mathfrak{S}(\mathrm{A})$ will become $\mathfrak{S}$). Moreover we will represent our models in Bloch form as introduced in Sect. 4, denoting by $\boldsymbol{l} = \{l_i\}$ and $\boldsymbol{\lambda} = \{\lambda_j\}$ the canonical basis of $\mathfrak{E}_{\mathbb{R}}$ and $\mathfrak{S}_{\mathbb{R}}$, respectively.

5 Toy-theory 1: the two-box model (extended Popescu-Rohrlich model)

The original model contains only states and effects, and has been already considered in Ref. [9] as a testing model for our present probabilistic framework. Here we will extend the model, by adding transformations in a consistent fashion.

5.1 Original model: the Popescu-Rohrlich boxes

The original model has been presented in Ref. [10]. It is locally made of a *b*ox which provides the probability rule for the output given the input. In the simplest situation, input and output are both binary. The probability rules are sketched in Fig. 3, and are given by[6]

$$\mathbb{P}_{\alpha\beta}(i|x) = \begin{cases} 1, & i = \alpha x \oplus \beta \\ 0 & \text{otherwise}, \end{cases} \qquad \alpha, \beta = 0, 1, \tag{58}$$

for the two possible outputs $i = 0, 1$ given the two possible inputs $x = 0, 1$.

The core of the original work are the correlated boxes in Fig. 4 defined by the joint probabilities $\mathbb{P}(ij|xy)$ consistent with no-signaling. As shown in Ref. [19], the complete set of these probabilities makes an eight dimensional polytope with 24 vertices.

[6] In Eq. (58) the symbol $\oplus$ denotes the addition modulo 2.

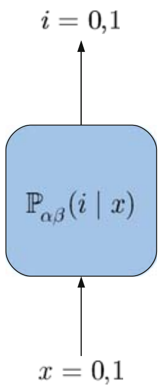


Fig. 3 The Popescu-Rohrlich box which provides the probability rules for the binary output given the binary input

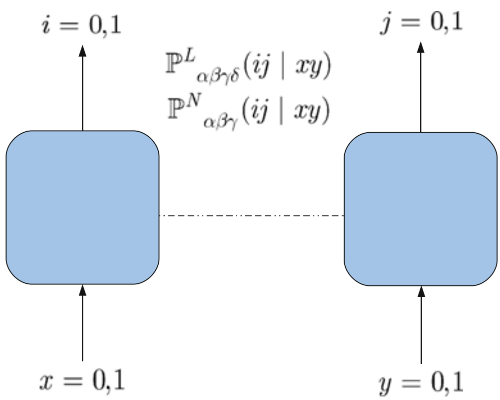


Fig. 4 The Popescu-Rohrlich correlated boxes which provide the probability rules for the binary output given the binary input

Among these 24 probability distributions we can identify two main classes, the local and the non-local boxes denoted as $\mathbb{P}^L_{\alpha\beta\gamma\delta}(ij|xy)$ and $\mathbb{P}^N_{\alpha\beta\gamma\delta}(ij|xy)$, and respectively given by:

$$\begin{aligned} \mathbb{P}^L_{\alpha\beta\gamma\delta}(ij|xy) &= \begin{cases} 1 & i = \alpha x \oplus \beta \\ & j = \gamma x \oplus \delta \\ 0 & \text{otherwise,} \end{cases} \\ \mathbb{P}^N_{\alpha\beta\gamma}(ij|xy) &= \begin{cases} \frac{1}{2} & i \oplus j = xy \oplus \alpha x \oplus \beta y \oplus \gamma \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \tag{59}$$

where $\alpha, \beta, \gamma, \delta \in \{0, 1\}$. The 16 local vertices $\mathbb{P}^L_{\alpha\beta\gamma\delta}(ij|xy)$ correspond to the factorization of the single box probability rules $\mathbb{P}_{\alpha\beta}(i|x)$, while the 8 non-local probability rules $\mathbb{P}^N_{\alpha\beta\gamma}(ij|xy)$ introduce the strongest correlations (corresponding to the maximal violation of the CHSH inequality) compatible with no-signaling.

In the following we will introduce the cones of states and effects, and we will extend the original model by introducing transformations. This will be done upon starting from a bipartite state considered as preparationally faithful.

5.2 Local sets of states and effects

According to the local box in Fig. 3 we can perform two possible tests, $\mathbb{A}^{(x)} = \{\mathscr{A}_0^{(x)}, \mathscr{A}_1^{(x)}\}$ with $x = 0, 1$. Correspondingly, we will denote the effects of the test $\mathbb{A}^{(x)}$ as $a_0^{(x)}, a_1^{(x)}$, with

$$a_0^{(0)} + a_1^{(0)} = a_0^{(1)} + a_1^{(1)} = e, \tag{60}$$

where e is the deterministic effect. There are only three independent local effects, whence $\dim(\mathfrak{E}_+) = \dim(\mathfrak{S}_+) = 3$. Clearly $\dim(\mathfrak{S}) = 2$, namely there are only two affinely independent states. The local convex set of states is the 2-dimensional *polytope* $\mathbb{P}^2$ given by the convex hull of the probability rules $\mathbb{P}_{\alpha\beta}(i|x)$ in Eq. (58). These are the vertices of $\mathfrak{S}$, namely the pure states of the model. In the following we will denote them by $\omega^{\alpha\beta}$.

It is convenient to represent the effects in a 3-dimensional vector space with the canonical coordinate along the z-axis corresponding to the deterministic effect e. Therefore a possible representation of the four effects in the two tests is

$$\boldsymbol{\lambda}(e) = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \ \boldsymbol{\lambda}(a_0^{(0)}) = \begin{bmatrix} \frac{1}{2} \\ -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \ \boldsymbol{\lambda}(a_1^{(0)}) = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix},$$
$$\boldsymbol{\lambda}(a_0^{(1)}) = \begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}, \ \boldsymbol{\lambda}(a_1^{(1)}) = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{2} \end{bmatrix}. \tag{61}$$

Correspondingly, according to the probability rule in Eq. (58), the four pure states will be represented as

$$\boldsymbol{l}(\omega^{00}) := \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \ \boldsymbol{l}(\omega^{11}) := \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \ \boldsymbol{l}(\omega^{01}) := \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \ \boldsymbol{l}(\omega^{10}) := \begin{bmatrix} 0 \\ -1 \\ 1 \end{bmatrix}. \tag{62}$$

One can easily verify the application of the states to the effects

$$\mathbb{P}_{\alpha\beta}(i|x) = \omega^{\alpha\beta}(a_i^{(x)}) \equiv \boldsymbol{l}(\omega^{\alpha\beta}) \cdot \boldsymbol{\lambda}(a_i^{(x)}) = \begin{cases} 1, & i = \alpha x \oplus \beta \\ 0 & \text{otherwise.} \end{cases} \tag{63}$$

Notice that the third coordinate along the axis of the cone $\mathfrak{S}_+$ is constantly equal to unit. Denoting by x, y, z the three components of vectors in both the Euclidean spaces $\mathfrak{S}_{\mathbb{R}}$ and $\mathfrak{E}_{\mathbb{R}}$, the 2-dimensional polytope $\mathbb{P}^2$ of states is (see the square at the top in Fig. 5)

$$\mathfrak{S} = \mathbb{P}^2 = \{\boldsymbol{l}(\omega) \mid |x| + |y| \leqslant 1\}, \tag{64}$$

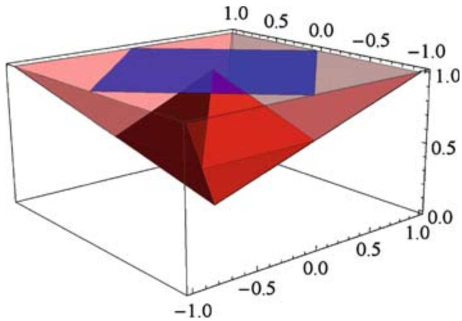


Fig. 5 The *square at the top* represents the set of states $\mathfrak{S}$. The *transparent cone* represents the dual cone of effects $\mathfrak{E}_+$. The *octahedron inside the transparent cone* represents the convex set of effects $\mathfrak{E}$ which is the $\mathfrak{E}_+$-truncation given by the condition $a \leqslant e$, where a is a generic effect and e the deterministic one

which is the convex hull of the vectors $\boldsymbol{l}(\omega^{\alpha\beta})$ ($\alpha = 0, 1,\ \beta = 0, 1$) corresponding to the vertices of $\mathfrak{S}$. Clearly the cone $\mathfrak{S}_+$, based on $\mathfrak{S}$, and its dual $\mathfrak{E}_+$ are given by

$$\mathfrak{S}_+ = \{\boldsymbol{l}(\omega) \,|\, |x| + |y| \leqslant z,\ z \geqslant 0\}, \quad \mathfrak{E}_+ = \{\boldsymbol{\lambda}(a) \,|\, |x| \leqslant z,\ |y| \leqslant z,\ z \geqslant 0\}. \tag{65}$$

Therefore the convex set of physical effects is

$$\mathfrak{E} = \left\{ \boldsymbol{\lambda}(a) \ \text{such that} \ \begin{Bmatrix} |x| \leqslant z,\ |y| \leqslant z, & z \in \left[0, \frac{1}{2}\right] \\ |x| \leqslant 1 - z,\ |y| \leqslant 1 - z, & z \in \left[\frac{1}{2}, 1\right] \end{Bmatrix} \right\}, \tag{66}$$

corresponding to the truncation of $\mathfrak{E}_+$ given by the order prescription $0 \leqslant a \leqslant e$.

5.3 The bipartite system and the faithful state

As mentioned, the joint probabilities $\mathbb{P}_{\alpha\beta\gamma}(ij|xy)$ form a table of 16 (2^4) entries, but only 8 of them are independent. Thus the bipartite convex set of states $\mathfrak{S}(\mathrm{AA})$ is the 8-dimensional polytope with the 24 vertices given in Eq. (59). These correspond to the pure bipartite states of the model: the 16 factorized states $\omega^{\alpha\beta} \otimes \omega^{\gamma\delta}$, plus the 8 non-local ones, which we will denote by $\Phi^{\alpha\beta\gamma}$. The whole set $\mathfrak{S}(\mathrm{AA})$ is then the convex hull of its vertices. A way to introduce the whole set of transformations[7] $\mathfrak{T}_+$ compatible with the cone of bipartite states $\mathfrak{S}_+(\mathrm{AA})$ is to assume the cone-isomorphism $\mathfrak{T}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{AA})$ induced by a preparationally faithful state Φ according to Postulate PFAITH. We can take one of the non-local vertices $\Phi^{\alpha\beta\gamma}$—say $\Phi = \Phi^{000}$— as a pure symmetric preparationally faithful state. First we have to check that, regarded as a matrix over effects, the state Φ is non singular, since a preparationally faithful state is also an isomorphic map between the cones $\mathfrak{S}_+$ and $\mathfrak{E}_+$. Indeed we have

$$\Phi = \sum_{ij} \Phi_{ij} \lambda_i \otimes \lambda_j \equiv \boldsymbol{\Phi} = \{\Phi_{ij}\} = \{\Phi(l_i, l_j)\}, \tag{67}$$

[7] We know that transformations are usually the completely positive maps but in this model as in the following ones we consider only two systems and then the transformations are two-positive maps by construction.

and from the rules on the right side of Eq. (59) we get the non singular matrix

$$\boldsymbol{\Phi} = \begin{bmatrix} \Phi(a_0^{(0)} - a_1^{(1)}, a_0^{(0)} - a_1^{(1)}) & \Phi(a_0^{(0)} - a_1^{(1)}, a_1^{(0)} - a_0^{(0)}) & \Phi(a_0^{(0)} - a_1^{(1)}, e) \\ \Phi(a_0^{(0)} - a_1^{(1)}, a_1^{(0)} - a_0^{(0)}) & \Phi(a_1^{(0)} - a_0^{(0)}, a_1^{(0)} - a_0^{(0)}) & \Phi(a_1^{(0)} - a_0^{(0)}, e) \\ \Phi(a_0^{(0)} - a_1^{(1)}, e) & \Phi(a_1^{(0)} - a_0^{(0)}, e) & \Phi(e, e) \end{bmatrix}$$
$$= \frac{1}{2}\begin{bmatrix} 1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 2 \end{bmatrix}. \tag{68}$$

The cone-isomorphism $\mathfrak{E}_+ \simeq \mathfrak{S}_+$ established by the map $\Phi(a, \cdot) = \omega_a$ is explicitly given by $\varphi_i := \Phi(l_i, \cdot)$, where the vectors φ_i are the images of the basis effects l_i under the map Φ. One also has

$$\Phi(a_0^{(0)}, \cdot) = \tfrac{1}{2}\omega^{00}, \quad \Phi(a_1^{(0)}, \cdot) = \tfrac{1}{2}\omega^{01}, \quad \Phi(a_0^{(1)}, \cdot) = \tfrac{1}{2}\omega^{10}, \quad \Phi(a_1^{(1)}, \cdot) = \tfrac{1}{2}\omega^{11}. \tag{69}$$

Notice that $\Phi(e, \cdot) = \chi$ has representative $\boldsymbol{l}(\chi) = \lambda_3$, namely it is the center of the square $\mathfrak{S}$.

The same arguments leading to the matrix representation of Φ^{000} can be iterated for each state $\Phi^{\alpha\beta\gamma}$, and all of them are faithful states of the theory.

5.4 Introducing transformations

As already stated the symmetric preparationally faithful state Φ^{000} induces the cone-isomorphism $\mathfrak{T}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{AA})$. The first step is to achieve from the isomorphism an explicit relation between elements in the two cones. Then by this relation the whole set $\mathfrak{T}_+$ could be generated from the cone of bipartite states $\mathfrak{S}_+(\mathrm{AA})$. Let $\mathscr{A}$ be a generic transformation in $\mathfrak{T}_+$. Then take the matrix representation of $\mathscr{A}$ induced by the relation

$$l_i \circ \mathscr{A} =: \sum_k A_{ik} l_k. \tag{70}$$

From the isomorphism $\mathfrak{T}_+ \simeq \mathfrak{S}_+(\mathrm{AA})$ we know that

$$\forall\ \Psi \in \mathfrak{S}_+(\mathrm{AA}) \qquad \exists!\ \mathscr{A}_\Psi \in \mathfrak{T}_+ \ \text{ such that } \ (\mathscr{I} \otimes \mathscr{A}_\Psi)\Phi = \Psi. \tag{71}$$

Matching the last two equations we have

$$(\mathscr{I} \otimes \mathscr{A}_\Psi)\Phi(l_i, l_j) = \Psi(l_i, l_j) \ \Rightarrow \Phi(l_i, l_j \circ \mathscr{A}_\Psi) = \Psi(l_i, l_j) \Rightarrow \sum_k \Phi_{ik} A_{ik} = \Psi_{ij}$$
$$\Rightarrow \boldsymbol{\Phi A}_\Psi^t = \boldsymbol{\Psi} \Rightarrow \boldsymbol{A}_\Psi = \boldsymbol{\Psi}^t \boldsymbol{\Phi}^{-1}. \tag{72}$$

It is sufficient to find the twenty-four extremal rays of $\mathfrak{T}_+$, namely the ones associated to the extremal rays of $\mathfrak{S}_+(\mathrm{AA})$, according to the cone-isomorphism $\mathfrak{T}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{AA})$.

First we achieve the transformations corresponding to the non-local vertices $\Phi^{\alpha\beta\gamma}$, namely the eight maps $\mathscr{D}^{\alpha\beta\gamma}$ such that

$$(\mathscr{I} \otimes \mathscr{D}^{\alpha\beta\gamma})\Phi = \Phi^{\alpha\beta\gamma} \qquad \alpha, \beta, \gamma = 0, 1. \tag{73}$$

From their representatives $\boldsymbol{D}^{\alpha\beta\gamma} = \Phi^{\alpha\beta\gamma}\Phi^{-1}$ it is easy to verify the identity $\mathbf{G}_{\mathrm{A}} = \{\mathscr{D}^{\alpha\beta\gamma}\}$, namely the maps $\mathscr{D}^{\alpha\beta\gamma}$ are the eight automorphisms of the local square of states $\mathfrak{S}$: $\mathscr{D}^{000}$, $\mathscr{D}^{111}$, $\mathscr{D}^{001}$, $\mathscr{D}^{110}$ performing respectively a 2π, $\pi/2$, π, $3\pi/2$, 2π-clockwise rotation around the axis of the cone $\mathfrak{S}_+$, while $\mathscr{D}^{100}$, $\mathscr{D}^{011}$, $\mathscr{D}^{010}$, $\mathscr{D}^{101}$ perform the four reflection-symmetries of the square of states. As a consequence of Postulate PFAITH (see Sect. 3.1) the transposed of the automorphisms must be still automorphisms, as it can be directly verified in this case. Moreover the application of the automorphisms to the faithful state Φ produces the eight pure bipartite states of $\mathfrak{S}(\mathrm{AA})$ which are all pure symmetric preparationally faithful. Finally, we can verify that the maximally chaotic state $\chi = \Phi_{000}(e, \cdot)$ is invariant under $\mathbf{G}_{\mathrm{A}}$, namely $\mathscr{D}^{\alpha\beta\gamma}\chi = \chi, \forall\ \mathscr{D}^{\alpha\beta\gamma} \in \mathbf{G}_{\mathrm{A}}$, as stated among the consequences of PFAITH in Sect. 3.1.

The other extremal elements of $\mathfrak{T}_+$ are the transformations associated to the sixteen pure states $\omega^{\alpha\beta} \otimes \omega^{\gamma\delta}$. From the explicit isomorphism in Eq. (72) we get the required sixteen transformations, given by the eight maps

$$\frac{1}{2}\begin{bmatrix} c & -c & c \\ 0 & 0 & 0 \\ 1 & -1 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} -c & c & c \\ 0 & 0 & 0 \\ -1 & 1 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} c & c & c \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}, \quad \frac{1}{2}\begin{bmatrix} -c & -c & c \\ 0 & 0 & 0 \\ -1 & -1 & 1 \end{bmatrix}, \quad c = \pm 1, \tag{74}$$

along with the eight maps given by inverting the first and the second rows. From these transformations plus the automorphisms in $\mathbf{G}_{\mathrm{A}} = \{\mathscr{D}^{\alpha\beta\gamma}\}$ it is possible to generate the extremal rays of the cone $\mathfrak{T}_+$ ($\mathsf{Erays}(\mathfrak{T}_+)$) and, by convex combinations, the whole set $\mathfrak{T}_+$.

As already mentioned in Observation 4.1, the extended Popescu-Rohrlich model has the following interesting property

Proposition 5.1 *The extremal transformations of the extended Popescu-Rohrlich model coincide with the extremal elements in the equivalences classes of all the extremal effects.*

Proof We know from Sect. 5.2 that $\mathrm{Extr}(\mathfrak{E}) = \{e, 0, a_0^{(0)}, a_0^{(1)}, a_1^{(0)} a_1^{(1)}\}$. According to the definition of $\mathsf{gen}(\mathfrak{E})$ given in Sect. 4, we can assume $\mathsf{gen}(\mathfrak{E}) = \{e, 0, a_0^{(0)}\}$. In fact acting on $a_0^{(0)}$ with $\mathbf{G}_{\mathrm{A}}$ the other extremal points of $\mathfrak{E}$ are achievable.

First we look for $\mathrm{Extr}(a_0^{(0)})$. From Proposition 4.1 we know that the Bloch representative $\boldsymbol{A} = \{A_{ij}\}$ of a transformation $\mathscr{A} \in a_0^{(0)}$ has $\boldsymbol{\lambda}(a_0^{(0)})$ as the last row, namely $A_{31} = 1/2$, $A_{32} = -1/2$ and $A_{33} = 1/2$. Moreover $\mathscr{A}$ must be positive and then $\mathscr{A}\omega^{\alpha\beta} \in \mathfrak{S}_+\ \forall\alpha\beta \in 0, 1$. Remembering the definitions of $\omega^{\alpha\beta}$ and $\mathfrak{S}_+$ the last conditions produce the four inequalities

$$
\begin{aligned}
&|A_{11} + A_{13}| + |A_{21} + A_{23}| \leqslant 1, \quad |-A_{12} + A_{13}| + |-A_{22} + A_{23}| \leqslant 1,\\
&|A_{12} + A_{13}| + |A_{22} + A_{23}| \leqslant 0, \quad |-A_{11} + A_{13}| + |-A_{21} + A_{23}| \leqslant 0. \quad (75)
\end{aligned}
$$

The third and the last bounds fix the equalities $A_{12} = -A_{13}$, $A_{21} = -A_{23}$, $A_{11} = A_{13}$ and $A_{21} = A_{23}$ making the positivity condition as simple as $|A_{11}| + |A_{21}| \leqslant 1/2$. The extremal matrices in the equivalence class $a_0^{(0)}$, namely $\mathrm{Extr}(a_0^{(0)})$, are the four maps

$$
\begin{bmatrix} c & c & c \\ 0 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 & 0 \\ c & c & c \\ \frac{1}{2} & -\frac{1}{2} & \frac{1}{2} \end{bmatrix}, \qquad c = \pm\frac{1}{2}. \quad (76)
$$

According to Proposition 4.3 the extremal elements in $a_0^{(1)}$, $a_1^{(0)}$ and $a_1^{(1)}$ follow from the application of $\mathbf{G}_\mathrm{A}$ to the matrices in Eq. (76). The result are exactly the sixteen maps associated to the sixteen pure states $\omega^{\alpha\beta} \otimes \omega^{\gamma\delta}$ by the cone-isomorphism $\mathfrak{S}_+(\mathrm{AA}) \simeq \mathfrak{T}_+(\mathrm{A})$.

Finally we consider the Bloch representatives of the deterministic transformations in e, whose last row is $\boldsymbol{\lambda}(e) = [0, 0, 1]$. A simple calculus, similar to the previous one, shows that $\mathrm{Extr}(e)$ is exactly the set of automorphisms $\mathbf{G}_\mathrm{A} = \{\mathscr{D}^{\alpha\beta\gamma}\}$ associated to the eight pure states $\Phi^{\alpha\beta\gamma}$ by the cone-isomorphism $\mathfrak{S}_+(\mathrm{AA}) \simeq \mathfrak{T}_+(\mathrm{A})$. □

5.5 Impossibility of teleportation

It is well known that the Popescu-Rohrlich model exhibits stronger nonlocality than Quantum Theory. For this reason one may argue that teleportation should be achievable. However, this is not the case, as we will see in the following. Consider for example the preparationally faithful state $\Phi^{000} = \Phi$ and the bilinear form F such that

$$
(F|_{23}\, |\Phi)_{12}\, |\Phi)_{34} = \alpha\, |\Phi)_{14}, \quad (77)
$$

for some $\alpha \in (0, 1]$. In order to satisfy Eq. (77) the matrix $\boldsymbol{F}$, which represents F in our Bloch basis, must be proportional to $\boldsymbol{\Phi}^{-1}$, namely

$$
F \propto (l_1 \otimes l_1) - (l_1 \otimes l_2) - (l_2 \otimes l_1) - (l_2 \otimes l_2) + (l_3 \otimes l_3). \quad (78)
$$

It is easy to verify that F is not a genuine bipartite effect. In fact, while the application of F to separable states always gives positive result

$$
F(\omega, \zeta) \geqslant 0 \qquad \forall\, \omega, \zeta \in \mathfrak{S}, \quad (79)
$$

exploring the application of F to bipartite states, we find

$$
\begin{aligned}
F(\Phi_{001}) &\propto \Phi_{001}(l_1, l_1) - \Phi_{001}(l_1, l_2)\\
&\quad - \Phi_{001}(l_2, l_1) - \Phi_{001}(l_2, l_2) + \Phi_{001}(l_3, l_3) = -1. \quad (80)
\end{aligned}
$$

This shows that Postulate FAITHE is not satisfied in this model and, according to Corollary 3.1, teleportation cannot be achieved in the extended Popescu-Rohrlich probabilistic theory.

Observation 5.1 *Notice that according to Proposition 3.2 the Popescu-Rohrlich theory does not admit a super-faithful state, which should achieve probabilistic teleportation.*

5.6 A theory without purification

Another fundamental quantum feature, the purifiability of all states, is not satisfied by the Popescu-Rohrlich model, namely Postulate PURIFY-1 does not hold. In fact the only pure bipartite states, apart from the sixteen factorized ones $\omega^{\alpha\beta} \otimes \omega^{\gamma\delta}$, are the eight maximally correlated states in Eq. (59) which are all purifications of the maximally chaotic state χ

$$\Phi_{\alpha\beta\gamma}(\cdot, e) = \Phi_{\alpha\beta\gamma}(e, \cdot) = \chi \qquad \forall \alpha, \beta, \gamma = 0, 1. \tag{81}$$

In conclusion there are too few pure bipartite states with respect to the infinite mixed states to be purified (the internal points of the square $\mathfrak{S}$). This will not be the case in the following class of models.

6 Toy-theory 2: the two-clock model

The Two-clock probabilistic models have a **clock** as local system, namely a system whose convex set of states is the disk $\mathbb{B}^2$. Many theories with such a local convex set of states can be generated: here we investigate their properties as possible probabilistic theories.

6.1 The self-dual local system

We can consider the model self-dual at the local system level. Therefore, in the usual representation, the cones of states and effects coincide

$$\mathfrak{S}_+ = \left\{ \boldsymbol{l}(\omega) | x^2 + y^2 \leqslant z^2, z \geqslant 0 \right\}, \quad \mathfrak{E}_+ = \left\{ \boldsymbol{\lambda}(a) | x^2 + y^2 \leqslant z^2, z \geqslant 0 \right\}, \tag{82}$$

namely the theory is (pointedly) self-dual at a single system level if we embed both the cones in the same Euclidean space $\mathbb{R}^3$. As usual, the deterministic effect in our canonical basis is given by the vector $\boldsymbol{\lambda}(e) = [0, 0, 1]$. The set of states $\mathfrak{S} \equiv \mathbb{B}^2$ is the base of the cone $\mathfrak{S}_+$ at $z = 1$, whereas the convex set of effects $\mathfrak{E}$ is the set of points of the cone $\mathfrak{E}_+$ satisfying $e - a \in \mathfrak{E}_+$, namely

$$\mathfrak{S} = \left\{ \boldsymbol{l}(\omega) | x^2 + y^2 = 1 \right\}, \;\; \mathfrak{E} = \left\{ \boldsymbol{\lambda}(a) | x^2 + y^2 \leqslant \min(z^2, (1 - z)^2), \;\; z \in [0, 1] \right\}. \tag{83}$$

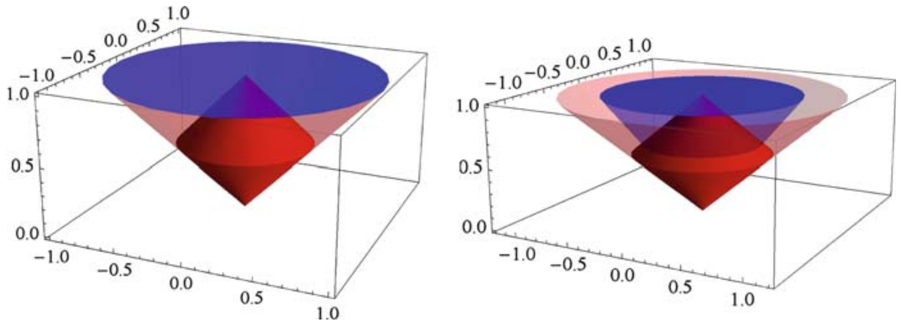


Fig. 6 *Left figure: the disk at the top* represents the set of states $\mathfrak{S}$. The *transparent cone* represents the cones $\mathfrak{S}_+$ and $\mathfrak{E}_+$. The *solid inside the transparent cone* represents the convex set of effects $\mathfrak{E}$ which is the $\mathfrak{E}_+$-truncation given by the condition $a \leqslant e$, e being the deterministic effect. *Right figure:* the same as in the left figure in the non self-dual case

Therefore, the convex set of effects $\mathfrak{E}$ is made of two truncated cones of height $\frac{1}{2}$ glued together at the basis, as in the left Fig. 6, with the two vertices given by the null and the deterministic effect, respectively.

6.2 The faithful state choice

We will now introduce the joint states of the model. Although the local cones do not identify uniquely the bipartite system, if the model has a faithful state its bipartite structure is tightly connected to the local one. We know that a faithful state must provide the automorphism $\mathfrak{S}_+ \simeq \mathfrak{E}_+$ between the local cones of states and effects, and this requirement narrows the possible choices for the faithful state itself. Let's introduce the bipartite functional

$$\Phi(a, b) = \sum_i \lambda_i(a)\lambda_i(b). \tag{84}$$

One can check that it is positive over the cone of effects, but also over its linear span. Φ can be taken as a pure preparationally faithful state. Indeed, Φ gets the cone-isomorphism $\mathfrak{S}_+ \simeq \mathfrak{E}_+$ via the map

$$\omega_a := \Phi(a, \cdot) = a, \tag{85}$$

in agreement with self-duality. Notice that, similarly to the Popescu-Rohrlich model, the deterministic effect corresponds to the state $\chi = \Phi(e, \cdot) = \lambda_3$ at the center of $\mathfrak{S}$.

In the two-box model we have generated $\mathfrak{T}_+$ from the given cone $\mathfrak{S}_+(\mathrm{AA})$ using the isomorphism $\mathfrak{S}_+(\mathrm{AA}) \simeq \mathfrak{T}_+(\mathrm{A})$ induced by the preparationally faithful state of the theory. Here we choose the cone of transformations $\mathfrak{T}_+$ and use the isomorphism induced by Φ to deduce the cone of bipartite states $\mathfrak{S}_+(\mathrm{AA})$. The explicit isomorphism is that of Eq. (72), namely $\boldsymbol{A}_\Psi = \boldsymbol{\Psi}^t\boldsymbol{\Phi}^{-1}$. Now each bipartite state has the same representative matrix of the corresponding transposed transformation because, in terms of the canonical basis, one has $\Phi = \sum_{i=1}^{3} \lambda_i \otimes \lambda_i$, that is $\boldsymbol{\Phi} = I_3$. Thus the isomorphism simply reads

$$\boldsymbol{\Psi} = \boldsymbol{A}_{\Psi}^{t}. \tag{86}$$

6.3 Transformations

We are left with the problem of searching among the positive maps, which are also two-positive: these will be the physical maps of our model. The extremal transformations $\mathsf{Erays}(\mathfrak{T}_+)$ are the maps sending $\mathrm{Extr}(\mathfrak{S})$ into an elliptical conic of $\mathsf{Erays}(\mathfrak{S}_+)$[8] and we will call these maps *elliptical-maps*. There exist three different kinds of elliptical-maps corresponding to the three different elliptical conics:

a. *Circular-maps.* In these case the map $\mathscr{A}$ sends $\mathrm{Extr}(\mathfrak{S})$ into a circle (which is a particular ellipse) and $\mathfrak{S}$ into a disk.
b. *Degenerate-maps.* An elliptical conic is said to be degenerate when the intersection between the cone and the plane is a line, namely the plane is tangent to the cone. In these case the map $\mathscr{A}$ sends $\mathrm{Extr}(\mathfrak{S})$ into an extremal ray of $\mathfrak{S}_+$.
c. *Strictly elliptical-maps.* In these case $\mathrm{Extr}(\mathfrak{S})$ is mapped into a true ellipse.

First notice that it is $\mathbf{G}_{\mathrm{A}} = \mathbf{O}(2)$, namely the local automorphisms of the model are the rotation $\mathscr{R}_\phi$ by an angle ϕ around the cone axis plus the reflections $\mathscr{S}_\phi$ through the axis at ϕ. The elliptical-maps correspond to the transformations $\mathscr{R}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{R}_\theta{}^t$, $\mathscr{S}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{S}_\theta{}^t$, $\mathscr{R}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{S}_\theta{}^t$ and $\mathscr{S}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{R}_\theta{}^t$ where $\mathscr{A}^\gamma$ is the transformation having the following Bloch representative

$$\boldsymbol{A}^\gamma = \begin{bmatrix} \gamma & 0 & (1-\gamma) \\ 0 & \sqrt{2\gamma - 1} & 0 \\ (1-\gamma) & 0 & \gamma \end{bmatrix}, \qquad \gamma \in \left[\tfrac{1}{2}, 1\right]. \tag{87}$$

For example the maps corresponding to $\mathscr{S}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{S}_\theta{}^t$ are

$$\begin{aligned} &\boldsymbol{S}_\phi \boldsymbol{A}^\gamma \boldsymbol{S}_\theta{}^t \\ &= \begin{bmatrix} \gamma\cos 2\theta\cos 2\phi + \sqrt{2\gamma-1}\sin 2\theta \sin 2\phi & \gamma\cos 2\theta \sin 2\phi - \sqrt{2\gamma-1}\sin 2\theta\cos 2\phi & (1-\gamma)\cos 2\theta \\ \gamma \sin 2\theta\cos 2\phi - \sqrt{2\gamma-1}\cos 2\theta \sin 2\phi & \gamma\sin 2\theta\sin 2\phi + \sqrt{2\gamma-1}\cos 2\theta\cos 2\phi & (1-\gamma)\sin 2\theta \\ (1-\gamma)\cos 2\phi & (1-\gamma)\sin 2\phi & \gamma \end{bmatrix} \\ &\phi, \theta \in (0, \pi], \ \gamma \in \left[\tfrac{1}{2}, 1\right], \end{aligned} \tag{88}$$

while the other three combinations, $\mathscr{R}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{R}_\theta{}^t$, $\mathscr{R}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{S}_\theta{}^t$ and $\mathscr{S}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{R}_\theta{}^t$, are exactly the same, apart from sign. The set of extremal rays $\mathsf{Erays}(\mathfrak{T}_+)$ is made of all the maps proportional to the above ones. According to the value of the parameter γ it is possible to identify the following three different kinds of maps.

[8] A conic section (or just a conic) is a curve obtained by intersecting a cone (more precisely a circular conical surface) with a plane.

a. For $\gamma = 1$ we achieve the circular-maps. These maps are exactly the rotations and the reflections, namely the local automorphisms $\mathbf{G}_\mathrm{A}$ of the model. Accordingly, the last row of their Bloch representatives is the deterministic effect $\boldsymbol{\lambda}(e) = [0, 0, 1]$.
b. For $\gamma = 1/2$ we achieve the degenerate-maps. Denoting by a_ϕ, with $\phi \in (0, 2\pi]$, the extremal effects lying on the circle at $z = 1/2$ in the left Fig. 6, these maps are exactly $\mathrm{Extr}(a_\phi)$ for $\phi \in (0, 2\pi]$. Consider for example the effect a_0 having representative $\boldsymbol{\lambda}(a_0) = [1/2, 0, 1/2]$. According to the Bloch representation, the extremal map in Eq. (87) (for $\gamma = 1/2$) has effect a_0. All the extremal maps having this effect, namely $\mathrm{Extr}(a_0)$, are achieved from the previous one by applying $\mathbf{G}_\mathrm{A}$ on the left. Moreover from Proposition 4.3 we know that $\{\mathrm{Extr}(a_\Phi),\ \phi \in (0, 2\pi]\} = \mathrm{Extr}(a_0) \circ \mathbf{G}_\mathrm{A}$.
c. For $\gamma \in (1/2, 1)$ we get the strictly elliptical-maps. These transformations belong to the non extremal effects (equivalence classes) whose Bloch representatives are the vectors $[(1-\gamma)\cos\phi, (1-\gamma)\sin\phi, \gamma]$, for $\gamma \in (1/2, 1)$ and $\phi \in (0, 2\pi]$.

Observation 6.1 *According to Observation 4.1, in this model, as in Quantum Theory, there exist extremal transformations corresponding to non extremal effects.*

6.4 The bipartite cone of states

We know that the isomorphism $\mathfrak{T}_+(\mathrm{A}) \simeq \mathfrak{S}_+(\mathrm{AA})$ induced by the chosen faithful state leads to the relation in Eq. (86) between bipartite states and physical maps. Then the same matrices representing the extremal maps $\mathscr{R}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{R}_\theta{}^t$, $\mathscr{S}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{S}_\theta{}^t$, $\mathscr{R}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{S}_\theta{}^t$ and $\mathscr{S}_\phi \circ \mathscr{A}^\gamma \circ \mathscr{R}_\theta{}^t$ represent all the pure bipartite states too (apart from normalization). For completeness we report explicitly the matrices representing the normalized states associated to the transformations $\mathscr{S}_\phi \circ \mathscr{A} \circ \mathscr{S}_\theta{}^t$

$$
\begin{aligned}
\Psi &= \frac{(\mathscr{S}_\theta \otimes \mathscr{S}_\phi \circ \mathscr{A}^\gamma)\Phi}{\Phi(e, a \circ \mathscr{S}_\theta)} \Rightarrow \boldsymbol{\Psi} = \frac{\boldsymbol{\Phi}(\boldsymbol{S}_\phi \boldsymbol{A}^\gamma \boldsymbol{S}_\theta{}^t)^t}{\boldsymbol{\lambda}(e)^t \boldsymbol{\Phi} \boldsymbol{\lambda}(a \circ \mathscr{S}_\theta)} = \frac{(\boldsymbol{S}_\phi \boldsymbol{A}^\gamma \boldsymbol{S}_\theta{}^t)^t}{\gamma} \\
&= \begin{bmatrix} \cos 2\theta \cos 2\phi + \frac{\sqrt{2\lambda-1}}{\lambda} \sin 2\theta \sin 2\phi & \sin 2\theta \cos 2\phi - \frac{\sqrt{2\lambda-1}}{\lambda} \cos 2\theta \sin 2\phi & \frac{(1-\lambda)}{\lambda} \cos 2\phi \\ \cos 2\theta \sin 2\phi - \frac{\sqrt{2\lambda-1}}{\lambda} \sin 2\theta \cos 2\phi & \sin 2\theta \sin 2\phi + \frac{\sqrt{2\lambda-1}}{\lambda} \cos 2\theta \cos 2\phi & \frac{(1-\lambda)}{\lambda} \sin 2\phi \\ \frac{(1-\lambda)}{\lambda} \cos 2\theta & \frac{(1-\lambda)}{\lambda} \sin 2\theta & 1 \end{bmatrix} \\
&\quad \phi, \theta \in (0, \pi],\ \gamma \in \left[\tfrac{1}{2}, 1\right].
\end{aligned}
\tag{89}
$$

Notice that the states associated to the degenerate-maps, are the *factorized bipartite pure states* given by $\boldsymbol{l}(\omega_\theta) \otimes \boldsymbol{l}(\omega_\phi)$, $\forall \omega_\theta, \omega_\phi \in \mathrm{Extr}(\mathfrak{S})$. The states corresponding to the circular and strictly elliptical-maps are the *non-local bipartite pure states* of the model. In particular, as will be investigated in a forthcoming publication, the states associated to local automorphisms $\mathbf{G}_\mathrm{A}$ achieve the Cirel'son's bound (see Ref. [20]) of the model.[9]

[9] The Cirel'son's bound of the two-clock model is the same of the Quantum Theory one, namely $2\sqrt{2}$.

6.5 Purifiability at the single system level

Differently from the Popescu-Rohrlich probabilistic model the two-clock model satisfies Postulate PURIFY-1 at the single system level as stated in the following Proposition.

Proposition 6.1 *In the two-clock model with the cones $\mathfrak{T}_+$ and $\mathfrak{S}_+(\mathrm{AA})$ respectively introduced in Secs. 6.3 and 6.4, any mixed local state has purification unique up to local automorphisms on the purifying system.*

Proof In the Bloch representation the marginalization on the first system of a bipartite state is simply the last column of its representative matrix. Consider then the pure bipartite states in Eq. (89). By taking the marginals on the first system we get the set of local states

$$\left\{ \Psi(\cdot, e) \equiv \boldsymbol{\Psi}\boldsymbol{\lambda}(e) = \begin{bmatrix} \frac{(1-\gamma)}{\gamma}\cos\phi \\ \frac{(1-\gamma)}{\gamma}\sin\phi \\ 1 \end{bmatrix}, \ \phi \in (0, 2\pi], \ \gamma \in \left[\tfrac{1}{2}, 1\right] \right\} = \mathfrak{S}, \quad (90)$$

which coincides with the whole set of states $\mathfrak{S}$, proving purifiabilility of the model. Uniqueness up to local automorphisms is easily verified. In fact, first notice that if Ψ is a purification of ω, i.e. $\Psi(\cdot, e) = \omega$, then also the states $(\mathscr{I} \otimes \mathscr{R}_\phi)\Psi$ and $(\mathscr{I} \otimes \mathscr{S}_\phi)\Psi$ are purifications of ω, because the last column of their representative matrices are the same of the $\boldsymbol{\Psi}$'s one. Then, suppose that there exists another purification of ω—say Ψ'— which is not connected to Ψ by a local automorphism acting on the second system. But, according to the pure bipartite states introduced in Sect. 6.4, there exist $\mathscr{D}_1, \mathscr{D}_2 \in \mathbf{G}_\mathrm{A}$ such that $\Psi' = (\mathscr{D}_1 \otimes \mathscr{D}_2)\Psi$ and then $\Psi' = (\mathscr{I} \otimes \mathscr{D}_2 \mathscr{D}_1{}^t)\Psi$ which contradicts the hypothesis. □

6.6 Exploring teleportation and purifiability

The probabilistic model introduced in this section does not allow teleportation, because the inverse of the preparationally faithful state is not a true bipartite effect. In fact considering the state $\Psi = (\mathscr{I} \otimes \mathscr{R}_\pi) \in \mathfrak{S}_+(\mathrm{AA})$ we get $\Phi^{-1}(\Psi) = -1$, which is negative. More precisely we get

$$\Phi^{-1}(\Psi) \leqslant 0 \qquad \forall \Psi = (\mathscr{I} \otimes \mathscr{R}_\phi)\Phi \ \text{ with } \ \phi \in [5\pi/6, 7\pi/6]. \quad (91)$$

Thus Postulate FAITHE does not hold in this model and according to Corollary 3.1 teleportation is not achievable. A good question is how the set $\mathfrak{T}_+$, and then $\mathfrak{S}_+(\mathrm{AA})$, has to be restricted in order to achieve a theory which allows teleportation preserving the purifiability. Indeed, reducing the set of transformations we also reduce $\mathfrak{S}_+(\mathrm{AA})$ and, by duality, the set of bipartite effects $\mathfrak{E}_+(\mathrm{AA})$ grows.

Observation 6.2 *One could try to get a theory with teleportation excluding some automorphisms from $\mathfrak{T}$. Indeed excluding rotations in $\mathbf{O}(2)$, the states Ψ in Eq. (91)*

are no longer states of the theory and Φ^{-1} *would be a true effect. On the other hand we cannot take reflections as the only physical automorphisms because* $\mathfrak{T}$ *is closed under composition and all rotations are achievable by composing two reflections. We could eventually reduce the set of physical automorphisms to* $\mathbf{SO}(2)$ *but obviously teleportation would be still impossible.*

Observation 6.3 *As in the two-box model, Proposition 3.2 ensures that also the two-clock model does not admit a super-faithful state.*

In the following we will use the abbreviation *purifiability of states*, to express the existence of purification of states, uniquely up to reversible channels on the purifying system. From the impossibility of achieving teleportation in the present model an interesting property for a general probabilistic theory follows.

Proposition 6.2 *In a probabilistic theory, purifiability of single system states does not imply purifiability at higher multipartite levels of the theory.*

Proof The proof of this statement is simply the counterexample given by the two-clock model constructed in this Section. In fact from Proposition 6.1 we know that the model allows a purification for every mixed local state, unique up to reversible channels on the purifying system. This means that uniqueness of purification holds at the single system level. On the other hand, according to Proposition 3.3, the same property at all the multipartite levels of the theory should imply the possibility of achieving probabilistic teleportation, which has been already excluded. □

6.7 A global feature from the local system structure

Here we observe a global feature of the two-clock probabilistic theories arising from the shape of the local cones.

Proposition 6.3 *It is impossible to construct a probabilistic theory having a disk as local set of states and a self-dual bipartite system at the same time.*

Proof The model constructed in this Section is self-dual at the single system level as represented in Fig. 6. From the local self-duality it follows that the bipartite system is self-dual in correspondence of its "local component", namely the factorized bipartite states $\omega_1 \otimes \omega_2$, with $\omega_1, \omega_2 \in \mathfrak{S}$, are proportional to bipartite effects ($a_1 \otimes a_2$ with $\omega_1 = \Phi(a_1, \cdot)$ and $\omega_2 = \Phi(a_2, \cdot)$). On the other hand, the bipartite system is not self-dual because of its "non-local component". Indeed not all the bipartite states associated through the faithful state $\Phi = \sum_{i=1}^{3} \lambda_i \otimes \lambda_i$ to the local automorphisms $\mathbf{G}_{\mathrm{A}}$ are proportional to bipartite effects. Regarding the states $\Psi_\phi = (\mathscr{I} \otimes \mathscr{R}_\phi)\Phi$ as bipartite functionals over $\mathfrak{S}(\mathrm{AA})$ we get, for example, $\Psi_{2\pi}(\Psi_\pi) < 0$, namely $\Psi_{2\pi}$ is not proportional to a bipartite effect.

Considering the bipartite states associated by the faithful state to the local automorphisms $\mathbf{G}_{\mathrm{A}}$, the only way to make them proportional to bipartite effects is to modify the faithful state of the theory. To achieve this goal the faithful state must be $\Phi' = \frac{1}{\sqrt{2}}(\lambda_1 \otimes \lambda_1 + \lambda_2 \otimes \lambda_2) + \lambda_3 \otimes \lambda_3$. We know that the faithful state induces also the isomorphism $\Phi'(a,\cdot)=\omega_a$ between the local cones of effects and states. Differently from the old faithful state Φ, the new one squeezes the local cone of states with respect to the cone of effects, as showed in the right Fig. 6, destroying the local self-duality of the model. Naturally a model without local self-duality cannot be seldual at the bipartite system level because of its factorized component. □

7 An hidden quantum model for the two-clock model: the rebit

In the class of probabilistic theories having a disk as local convex set of states a special case is that of the *equatorial qubit*. In fact, the convex set of states for the *qubit* is the 3-dimensional ball known as *Bloch sphere*, and the clock corresponds to the qubit in the equatorial plane. This model is also called **rebit**, where "re" stays for real, and corresponds to Quantum Theory on a two-dimensional real Hilbert space. The peculiarity of the rebit model is that it violates *local observability*.

7.1 Local states and effects

Consider as usual the canonical basis $\boldsymbol{l}=\{l_i\}$ and $\boldsymbol{\lambda}=\{\lambda_i\}$ with $i=1,2,3$ for $\mathfrak{E}_{\mathbb{R}}$ and $\mathfrak{S}_{\mathbb{R}}$ embedded into $\mathbb{R}^3$ as Euclidean spaces. Inspired by the well known qubit model, upon defining the operator vector $\sigma=[\sigma_z,\sigma_x,I]$, and introducing the canonical orthonormal basis $\{u_j\}$ for $\mathbb{R}^3$, we define the following bijective map

$$\Upsilon : r\in\mathbb{R}^3 \leftrightarrow \Upsilon(r)\in \mathsf{Her}(\mathbb{R}^2),\quad \begin{cases}\Upsilon(r)=\frac{1}{\sqrt{2}}r\cdot\sigma,\\ \Upsilon^{-1}(A)=\frac{1}{\sqrt{2}}\operatorname{Tr}[A\sigma]\cdot u,\end{cases} \tag{92}$$

where u is the vector having the $\mathbb{R}^3$ basis vectors $\{u_i\}$ as components. We get the pairing relation[10]

$$\Upsilon(r)\bullet\Upsilon(s):=\operatorname{Tr}[\Upsilon(r)\Upsilon(s)],\quad \operatorname{Tr}[AB]=\Upsilon^{-1}(A)\cdot\Upsilon^{-1}(B). \tag{93}$$

The symbol $\bullet$ denotes a "scalar product" between elements in $\mathsf{Her}(\mathbb{R}^2)$ as defined in the last equation, and it is easy to verify that $\Upsilon(r)\bullet\Upsilon(s)=r\cdot s$, $\forall r,s\in\mathbb{R}^3$. In terms of the canonical basis one has

$$\Upsilon(u_i)=\Upsilon(l_i)=\Upsilon(\lambda_i)=\frac{1}{\sqrt{2}}\sigma_i,\quad (l_j,\lambda_i)=\frac{1}{2}\Upsilon^{-1}(\sigma_i)\cdot\Upsilon^{-1}(\sigma_j). \tag{94}$$

Specializing the map to states and effects of a clock we have the states and effects of the rebit (the hidden quantum model)

$$\omega\in\mathfrak{S},\ \rho=\frac{1}{\sqrt{2}}\Upsilon(\boldsymbol{l}(\omega))\in\mathsf{St}(\mathbb{R}^2),\quad a\in\mathfrak{E},\ A=\sqrt{2}\Upsilon(\boldsymbol{\lambda}(a))\in\mathsf{Lin}_+(\mathbb{R}^2), \tag{95}$$

10 One has: $\Upsilon^{-1}(A)\cdot\Upsilon^{-1}(B)=\frac{1}{2}\operatorname{Tr}[A\sigma]\cdot\operatorname{Tr}[B\sigma]=\operatorname{Tr}[(A\otimes B)\frac{1}{2}\sum_i\sigma_i\otimes\sigma_i]=\operatorname{Tr}[AB^t]-\operatorname{Tr}[(A\otimes B)\sigma_y\otimes\sigma_y]=\operatorname{Tr}[AB]$, where we have subtracted the component concerning σ_y.

with Born rule

$$\mathrm{Tr}[A\rho] = \Upsilon(\boldsymbol{l}(\omega)) \bullet \Upsilon(\boldsymbol{\lambda}(a)) \equiv (a, \omega), \tag{96}$$

$\mathsf{St}(\mathbb{R}^2)$ denoting the set of symmetric real matrices with unit trace. Notice that $\rho = \frac{1}{2}\left(I + \hat{\boldsymbol{l}}(\omega) \cdot \sigma\right)$, where $\hat{\boldsymbol{l}}(\omega)$ is the *Bloch vector* representing the point in the disk of states $\mathfrak{S}$. The extension of the map to tensor product is given by the "commutation rule" $\Upsilon\otimes = \otimes\Upsilon$, namely

$$\Upsilon(r \otimes s) := \Upsilon(r) \otimes \Upsilon(s), \quad \Upsilon^{-1}(A \otimes B) = \Upsilon^{-1}(A) \otimes \Upsilon^{-1}(B). \tag{97}$$

In the following we will use the abbreviate notation $\Upsilon(\omega) := \Upsilon(\boldsymbol{l}(\omega))$ for states and $\Upsilon(a) := \Upsilon(\boldsymbol{\lambda}(a))$ for effects.

7.2 The bipartite system: states and transformations

The faithful state is the bipartite functional Φ such that $\Phi(l_i, l_j) = \delta_{ij}$, whence the corresponding operator is given by

$$\tfrac{1}{2}\Upsilon(\Phi) = \tfrac{1}{2}\sum_{i=1}^{3} \Upsilon(\lambda_i) \otimes \Upsilon(\boldsymbol{\lambda}_i) = \tfrac{1}{4}(I \otimes I + \sigma_x \otimes \sigma_x + \sigma_z \otimes \sigma_z), \tag{98}$$

which is an Hermitian (non positive) operator with unit trace. Notice that such operator differs from the quantum maximally entangled state

$$\tfrac{1}{2}|I\rangle\!\rangle\langle\!\langle I| = \tfrac{1}{4}(I \otimes I + \sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y + \sigma_z \otimes \sigma_z), \tag{99}$$

by the term $\frac{1}{4}\sigma_y \otimes \sigma_y \notin \mathsf{Lin}(\mathbb{R}^2) \otimes \mathsf{Lin}(\mathbb{R}^2)$. The term $\sigma_y \otimes \sigma_y \in \mathsf{Lin}(\mathbb{R}^4)$ corresponds to the null linear form Ξ over $\mathbb{R}^3 \otimes \mathbb{R}^3$ given by

$$\Xi(R) = \mathrm{Tr}[\sigma_y \otimes \sigma_y \Upsilon(R)] = 0, \qquad \forall R \in \mathbb{R}^3 \otimes \mathbb{R}^3. \tag{100}$$

Notice that the transposition acts as the identity map over $\Upsilon(\mathbb{R}^3)$, since transposition leaves σ_x, σ_z and I invariant, whence $\Upsilon^{-1}[\Upsilon(a)^t] = \mathscr{I}(a)$. Using this identity one can also see that the maximally entangled state is another equivalent representation of the faithful state Φ, since $\forall r, s \in \mathbb{R}^3$ one has

$$\tfrac{1}{2}\langle\!\langle I|\Upsilon(\boldsymbol{\lambda}(a)) \otimes \Upsilon(\boldsymbol{\lambda}(b))|I\rangle\!\rangle = \tfrac{1}{2}\,\mathrm{Tr}[\Upsilon(\boldsymbol{\lambda}(a))\Upsilon(\boldsymbol{\lambda}(b))^t] = \boldsymbol{\lambda}(a) \cdot \boldsymbol{\lambda}(b) = \Phi(a, b) \tag{101}$$

(transposition works as the identity over σ_x, σ_z and I).

Let's now represent maps in the hidden quantum model. A generic bipartite state is represented as

$$\Psi = \sum_{ij} \Psi_{ij} \lambda_i \otimes \lambda_j, \quad \Psi_{ij} = \Psi(l_i, l_j), \tag{102}$$

and the local action of the transformation $\mathscr{A}$ is given by

$$\begin{aligned} (\mathscr{A} \otimes \mathscr{I})\Psi(l_i, l_j) &= \Psi(l_i \circ \mathscr{A}, l_j) = \sum_k A_{ik} \Psi(l_k, l_j) \\ &= \sum_{nklm} A_{nk} \Psi_{lm} \lambda_n(l_i) l_k(\lambda_l) \lambda_m(l_j) = \tfrac{1}{2} \operatorname{Tr}[\Upsilon(\mathscr{A}) \star \Upsilon(\Psi) \sigma_i \otimes \sigma_j], \end{aligned} \tag{103}$$

where

$$\Upsilon(\mathscr{A}) := \tfrac{1}{2} \sum_{nk} A_{nk} \sigma_n \otimes \sigma_k, \qquad \Upsilon(\Psi) := \tfrac{1}{2} \sum_{lm} \Psi_{lm} \sigma_l \otimes \sigma_m, \tag{104}$$

and

$$A \star B = \operatorname{Tr}_2[(A \otimes I)(I \otimes B)]. \tag{105}$$

The algebra of transformations allows a representation as operator algebra over $\mathsf{Her}(\mathbb{R}^2)$ and denoting by $\mathcal{A}$ ($\tilde{\mathcal{A}}$) and $\mathcal{I}$ the operators corresponding respectively to $\mathscr{A}$ ($\mathscr{A}'$) and $\mathscr{I}$ one has

$$(\mathscr{A} \otimes \mathscr{I})\Psi(l_i, l_j) = \tfrac{1}{2} \operatorname{Tr}[(\mathcal{A} \otimes \mathcal{I})\Upsilon(\Psi)\sigma_i \otimes \sigma_j] = \tfrac{1}{2} \operatorname{Tr}[\Upsilon(\Psi)\tilde{\mathcal{A}}(\sigma_i) \otimes \sigma_j], \tag{106}$$

whence

$$\Upsilon[(\mathscr{A} \otimes \mathscr{I})\Psi] = \Upsilon(\mathscr{A}) \star \Upsilon(\Psi) = (\mathcal{A} \otimes \mathcal{I})\Upsilon(\Psi) = \Upsilon(\Psi)(\tilde{\mathcal{A}} \otimes \mathcal{I}). \tag{107}$$

Now we have to choose the transformations of the model. In the previous two clocks models $\mathfrak{S}$ was a 2-dimensional convex set, whence $\dim(\mathfrak{S}_\mathbb{R}) = 3$ and $\dim(\mathfrak{T}_\mathbb{R}) = 9$. The set $\mathfrak{T}^q_\mathbb{R}$ for the qubit model is the linear Span of the quantum operations $\sigma_i \cdot \sigma_j$, for $i, j = 1, 2, 3, 4$ (we denote $\sigma_4 := \sigma_y$) and then

$$\mathfrak{T}^q_\mathbb{R} = \mathsf{Span}\{\mathscr{A}_{11}, \mathscr{A}_{22}, \mathscr{A}_{33}, \mathscr{A}_{44}, \mathscr{A}_{12}, \mathscr{A}_{13}, \mathscr{A}_{23}, \mathscr{A}_{14}, \mathscr{A}_{24}, \mathscr{A}_{34}\} \tag{108}$$

where

$$\mathscr{A}_{ij}\omega := \Upsilon^{-1}[\sigma_i \Upsilon(\omega)\sigma_j], \qquad \forall \omega \in \mathfrak{S}. \tag{109}$$

Notice that $\dim(\mathfrak{T}^q_{\mathbb{R}}) = 10$[11]. Here we are considering the equatorial qubit (rebit) and the space $\mathfrak{T}_{\mathbb{R}} \equiv \mathsf{Lin}(\mathfrak{E}_{\mathbb{R}}) = \mathsf{Lin}(\mathbb{R}^3)$ of linear maps over $\mathbb{R}^3$ can be obtained from the one in Eq. (108) as follows[12]

$$\mathfrak{T}^r_{\mathbb{R}} = \mathsf{Span}\,\{\mathscr{A}_{11}, \mathscr{A}_{22}, \mathscr{A}_{33}, \mathscr{A}_{44}, \Re\mathscr{A}_{12}, \Re\mathscr{A}_{13}, \Re\mathscr{A}_{23}, \Im\mathscr{A}_{14}, \Im\mathscr{A}_{24}, \Im\mathscr{A}_{34}\} \quad (110)$$

with $\mathscr{A}_{ij}$ as in Eq. (109).

We know that the automorphisms of the convex set of states $\mathfrak{S}$ are given by the rotations $\mathscr{R}_\phi,\ \ \phi \in [0, 2\pi)$ along with the reflections $\mathscr{S}_\phi,\ \ \phi \in [0, \pi)$ through the axis at ϕ. Taking the physical maps as in Eq. (110) we get all rotations and reflections of the disk of states. In fact the quantum operations achieve the automorphisms of the qubit system namely the rotations in $\mathbf{SO}(3)$. On the other hand the rotations of a sphere include not only the rotations of its equatorial disk but also its reflections. Therefore, $\mathbf{G}_{\mathrm{A}} = \mathbf{O}(2)$.

7.3 Ghosts

As already mentioned the set of transformations of the hidden quantum model (inherited from the qubit quantum operations) has dimension 10. On the other hand not all the matrices representing the 10 independent quantum operations are linearly independent when applied to the rebit. In fact the completely positive maps

$$\sigma_x \cdot \sigma_x + \sigma_z \cdot \sigma_z - I \cdot I, \qquad \sigma_y \cdot \sigma_y, \quad (111)$$

are not distinguishable by their local action over a rebit. As can be easily verified, the matrices representing the quantum operations in Eq. (111), which are locally distinguishable on a qubit, become the same on a re-bit. Clearly, by identification of locally indistinguishable transformations (namely taking the space of transformations with dimension 9), the local observability principle is satisfied. This is not the case if the space of transformations has dimension 10. Here indeed there exist two transformations indistinguishable by local tests but discriminable by bipartite measurements.

7.4 Bipartite effects and teleportation

In Eq. (93) we have defined the product $\Upsilon(r) \bullet \Upsilon(s) := \mathrm{Tr}[\Upsilon(r)\Upsilon(s)],\ \forall r, s \in \mathbb{R}^3$, from which the local states effects pairing $\Upsilon(\omega) \bullet \Upsilon(a) \equiv (a, \omega)$. We can coherently

[11] The qubit model is based on the 2-dimensional Hilbert space H and $\dim(\mathfrak{S}) = 4$ where $\mathfrak{S} = \mathsf{S}(\mathsf{H})$ is the states space. According to the Choi-Jamiolkowski isomorphism, in the 16-dimensional linear space $\mathsf{Lin}(\mathsf{H} \otimes \mathsf{H})$ we take only the operator corresponding to completely positive maps and we get $\dim(\mathfrak{T}^q_{\mathbb{R}}) = 10$ (the only Hermitian matrices are allowed).

[12] The symbols $\Re$ and $\Im$ stay respectively for Real and Imaginary part.

extend the product $\bullet$ as follows

$$\Upsilon(R) \bullet \Upsilon(S) = \tfrac{1}{2} \sum_{i,j} \mathrm{Tr}[\Upsilon(R) \star \Upsilon(S) \sigma_i \otimes \sigma_j] \qquad \forall R, S \in \mathbb{R}^3 \otimes \mathbb{R}^3, \quad (112)$$

to represent the pairing relation between bipartite states and effects as

$$\Upsilon(E) \bullet \Upsilon(\Psi) = (E, \Psi). \quad (113)$$

Proposition 7.1 *The rebit model does not allow probabilistic teleportation, nor a superfaithful state.*

Proof Let's first take the generalized effect corresponding to the inverse matrix of Φ (i.e. which would achieve teleportation), and let see if it is a true effect. The matrix multiplication between two (considering Φ^{-1} as a map) must be as follows

$$\delta_{ij} = \tfrac{1}{2} \mathrm{Tr}[\Upsilon(\Phi^{-1}) \star \Upsilon(\Phi) \sigma_i \otimes \sigma_j], \quad (114)$$

and taking $F^{\Phi} = \alpha \Phi^{-1}$ we get

$$(F^{\Phi}, \Phi) = \alpha(\Phi^{-1}, \Phi) = \alpha \Upsilon(\Phi^{-1}) \bullet \Upsilon(\Phi) = 3\alpha, \quad (115)$$

whence $\alpha = 1/3$, and one would have probability of successful teleportation

$$F^{\Phi} \omega \Phi(e) = (F^{\Phi}, e)(\omega, \Phi) = \alpha \, \mathrm{Tr}[(\Upsilon(\Phi^{-1}) \otimes \tfrac{1}{\sqrt{2}} I)(\Upsilon(\omega) \otimes \Upsilon(\Phi))] = \tfrac{1}{3}. \quad (116)$$

However, F^{Φ} is not a true effect. Consider the state Ψ given by

$$\Psi = (\mathscr{A}_{44} \otimes \mathscr{I}) \Phi = \Upsilon^{-1}[(\sigma_4 \otimes \mathcal{I}) \Upsilon(\Phi)(\sigma_4 \otimes \mathcal{I})], \quad (117)$$

where $\sigma_4 \cdot \sigma_4$ is a completely positive map, whence a transformation of the model. Explicitly

$$\begin{aligned} \tfrac{1}{2} \Upsilon(\Psi) &= \tfrac{1}{2} \Upsilon(\mathscr{A}_{44}) \star \Upsilon(\Phi) = \tfrac{1}{2} (\sigma_4 \otimes \mathcal{I}) \Upsilon(\Phi)(\sigma_4 \otimes \mathcal{I}) \\ &= \tfrac{1}{4} (I \otimes I - \sigma_x \otimes \sigma_x - \sigma_z \otimes \sigma_z). \end{aligned} \quad (118)$$

If we take the scalar product (F^{Φ}, Ψ), we have

$$(F^{\Phi}, \Psi) = \Upsilon(F^{\Phi}) \bullet \Upsilon(\Psi) = \tfrac{1}{3} \Upsilon(\Phi^{-1}) \bullet \Upsilon(\Psi) = -1, \quad (119)$$

namely a negative value, whence F^{Φ} is not a bipartite effect. Thus, Postulate FAITHE is not satisfied, and according to Corollary 3.1 teleportation is not achievable. Moreover, from Proposition 3.2, the rebit probabilistic theory does not allow a super-faithful state. □

7.5 Purifiability

It is well known that the Quantum Theory of real matrices satisfies Postulate PURIFY-1. For each local state $\Upsilon(\omega) = \rho_\omega = (I + \hat{\boldsymbol{l}}(\omega) \cdot \sigma)/2$ of the rebit system, we find a pure bipartite state $|\,{\rho_\omega}^{1/2}\rangle\!\rangle$ which purifies it.

The bipartite state $\Upsilon(\Psi)$ corresponding to $|\,{\rho_\omega}^{1/2}\rangle\!\rangle$ is given by the relation

$$\tfrac{1}{2}\Upsilon(\Psi) = |\,{\rho_\omega}^{1/2}\rangle\!\rangle\langle\!\langle{\rho_\omega}^{1/2}| \quad \text{with} \quad \mathrm{Tr}_2\left[|\,{\rho_\omega}^{1/2}\rangle\!\rangle\langle\!\langle{\rho_\omega}^{1/2}|\right] = \rho_\omega. \tag{120}$$

All the purifications of a state are connected by local automorphisms on the purifying system, that is $(e|_2\,(\mathscr{I} \otimes \mathscr{D})\,|\Psi) = |\omega)_1 \;\forall \mathscr{D} \in \mathbf{G}_\mathrm{A}$, or in quantum notation,

$$\mathrm{Tr}_2\left[(\mathcal{I} \otimes \mathcal{D})|\,{\rho_\omega}^{1/2}\rangle\!\rangle\langle\!\langle{\rho_\omega}^{1/2}|(\mathcal{I} \otimes \tilde{\mathcal{D}})\right] = \rho_\omega. \tag{121}$$

In the last equation we have used the relation $\mathcal{D}\tilde{\mathcal{D}} = \mathcal{I}$.

We have already shown that FAITHE is not satisfied. Therefore, from Proposition 3.3, the uniqueness of purification, up to reversible channels, at all the multipartite levels, is not satisfied.

8 Toy-theory 3: the two-spin-factor model

The convex set of states of the clock is the disk $\mathfrak{S} = \mathbb{B}^2$, whereas for the qubit one has $\mathfrak{S} = \mathbb{B}^3$. It seems then interesting to investigate probabilistic theories with $\mathfrak{S} = \mathbb{B}^n$. The local system of these theories is denoted **(n)spin-factor**. Naturally, as noticed for the two-clock model, many probabilistic theories may have the same (n)spin-factor as local system.

8.1 The self-dual (n)spin-factor, its states and effects

Consider the self-dual (n)spin-factor and denote as usual by $\boldsymbol{l} = \{l_i\}$ and $\boldsymbol{\lambda} = \{\lambda_j\}$, with $i, j = 1, \ldots, n+1$, the canonical basis for $\mathfrak{S}_\mathbb{R}$ and $\mathfrak{E}_\mathbb{R}$. The cones of states and effects coincide, whence

$$\mathfrak{S}_+ = \left\{\boldsymbol{l}(\omega)|x_1^2 + \ldots + x_n^2 \leqslant x_{n+1}^2\right\}, \quad \mathfrak{E}_+ = \left\{\boldsymbol{\lambda}(a)|x_1^2 + \ldots + x_n^2 \leqslant x_{n+1}^2\right\}. \tag{122}$$

Naturally the set of states is the section of the cone at $x_{n+1} = 1$, while its truncation, from the order relation $0 \leqslant a \leqslant e$, gives the set of effects

$$\begin{aligned} \mathfrak{S} &= \left\{\boldsymbol{l}(\omega) \mid x_1^2 + \ldots + x_n^2 \leqslant 1\right\}, \\ \mathfrak{E} &= \left\{\boldsymbol{\lambda}(a) \mid \; x_1^2 + \ldots + x_n^2 \leqslant \min\left(x_{n+1}^2, (1 - x_{n+1})^2\right), \; x_{n+1} \in [0, 1]\right\}. \end{aligned} \tag{123}$$

8.2 What is special about the (3)spin-factor?

As for the clocks—the (2)spin-factors—the probabilistic theory is defined only at the single system level. Therefore we need to extend the theory at the bipartite level. We reach this goal this by assuming as faithful state the $(n+1)$-dimensional generalization of the one given in Eq. (84), namely the bipartite functional

$$\Phi = \sum_{i=1}^{n+1} \lambda_i \otimes \lambda_i. \tag{124}$$

Being represented by the identical matrix $\boldsymbol{\Phi} = \boldsymbol{I}_{n+1}$, the state Φ realizes the cone-isomorphism $\mathfrak{S}_+ \simeq \mathfrak{E}_+$ via the map $\omega_a := \Phi(a, \cdot) = a$. In our probabilistic framework, from the isomorphism $\mathfrak{S}_+(\mathrm{AA}) \simeq \mathfrak{T}_+$ given by

$$\Psi = (\mathscr{I} \otimes \mathscr{A})\Phi \Rightarrow \boldsymbol{\Psi} = \boldsymbol{A}^t, \tag{125}$$

the cone of bipartite states $\mathfrak{S}_+(\mathrm{AA})$ can be generated from the set $\mathfrak{T}_+$ of two-positive maps (the transformations of our model), while the bipartite set of effects $\mathfrak{E}_+(\mathrm{AA})$ follows by duality from $\mathfrak{S}_+(\mathrm{AA})$.

The analysis of the spin-factors probabilistic model is extremely technical, and we will only report the main interesting result. First notice that for an (n)spin-factor, since the set of states is $\mathfrak{S} = \mathbb{B}^2$, one has $\mathbf{G}_\mathrm{A} \subset \mathbf{O}(n)$. Therefore the following proposition holds

Proposition 8.1 *Consider a probabilistic theory having an* (n)*spin-factor as local system, for each* $n \in \mathbb{N}$*. If* $\mathbf{G}_\mathrm{A} = \mathbf{O}(n)$*, then Postulate FAITHE is not satisfied.* □

Proof If $\mathbf{G}_\mathrm{A} = \mathbf{O}(n)$, namely if all the elements in the group $\mathbf{O}(n)$ are transformations of the theory, then it is always possible to find a bipartite state $\Psi \in \mathfrak{S}(\mathrm{AA})$ such that $F(\Psi) < 0$, where $F = \alpha\Phi^{-1}$ is the bipartite functional inverting the faithful state Φ (notice that from Eq. (124) $\boldsymbol{\Phi}^{-1} = \boldsymbol{I}_{n+1}$). In fact, consider the automorphism $\mathscr{D} \in \mathbf{O}(n)$ reversing the direction of every vector. The $(n+1) \times (n+1)$ matrix $\boldsymbol{D}$ representing $\mathscr{D}$ in our basis is the diagonal matrix with $D_{ii} = -1$ for $i = 1, \dots, n$ and $D_{n+1,n+1} = 1$. Therefore the state $\Psi = (\mathscr{I} \otimes \mathscr{D})\Phi = -\sum_{i=1}^{n}(\lambda_i \otimes \lambda_i) + \lambda_{n+1} \otimes \lambda_{n+1}$ achieves

$$F(\Psi) = -\sum_{i,j=1}^{n} l_i(\lambda_j) \otimes l_i(\lambda_j) + l_{n+1}(\lambda_{n+1}) \otimes l_{n+1}(\lambda_{n+1}) < 0 \quad \forall n \geqslant 2. \tag{126}$$

In general the automorphism $\mathscr{D}$ is a combination of reflections and rotations and it is not the only combination achieving a state Ψ with $F(\Psi) < 0$. It is possible to reduce the set of physical automorphisms from $\mathbf{O}(n)$ to its subgroup $\mathbf{SO}(n)$, that is the component connected to the identical transformation. However, the following proposition holds:

Proposition 8.2 *Consider a probabilistic theory having as local system an* (n)*spin-factor, for each* $n \in \mathbb{N}$*, as local system. If* $\mathbf{G}_{\mathrm{A}} = \mathbf{SO}(n)$ *then Postulate FAITHE is still violated.*

Proof It is easy to see that

$$\forall n \neq 3 \qquad \exists \mathscr{D} \in \mathbf{SO}(n) \;\; \text{such that} \;\; (F|\,(\mathscr{I} \otimes \mathscr{D})\,|\Phi) < 0, \tag{127}$$

and then $\forall n \neq 3$ Postulate FAITHE fails. For even $n \geqslant 2$ the situation is the same of Proposition 8.1 because the automorphism $\mathscr{D}$ reversing the direction of every vector is a rotation. For odd n, in order to achieve a Ψ such that $F(\Psi) < 0$, it is sufficient to take the automorphism $\mathscr{D}$ corresponding to the rotation of the n-dimensional ball around the n-th axis. The representative of $\mathscr{D}$ is the diagonal matrix with $D_{ii} = -1$ for $i = 1, \ldots, n-1$ and $D_{nn} = D_{n+1n+1} = 1$. Therefore the state $\Psi = (\mathscr{I} \otimes \mathscr{D})\Phi = -\sum_{i=1}^{n-1}(\lambda_i \otimes \lambda_i) + \lambda_n \otimes \lambda_n + \lambda_{n+1} \otimes \lambda_{n+1}$ achieves $F(\Psi) < 0$ for each odd $n \geqslant 5$. □

The last two Propositions show that, among the probabilistic theories having as local system an (n)spin-factor with $\mathbf{SO}(n) \subset \mathbf{G}_{\mathrm{A}}$, it is possible to satisfy Postulate FAITHE if and only if $n = 3$. Therefore, according to Corollary 3.1 and Proposition 3.3 teleportation and uniqueness (modulo local automorphisms) of purification at all levels can be satisfied. This is not surprising because the *qubit* is exactly the hidden quantum model (in the sense of Sect. 7) of the (3)spin-factor probabilistic theory having $\mathbf{SO}(3)$ as physical automorphisms.

9 Toy-theory 4: the classical probabilistic model

A probabilistic theory is said to be classical if and only if its local set of states $\mathfrak{S}$ is a *simplex*. Including these theories in our probabilistic framework we can easily show how some fundamental features of the classical theories arise from the simplex nature of $\mathfrak{S}$. Differently from the previous models the classical ones can be easily investigated for a general dimension.

9.1 Probability simplex representation

Consider a simplex set of states $\mathfrak{S}$ with $\dim(\mathfrak{S}) = n$ and denote as usual by $\boldsymbol{l} = \{l_i\}$ and $\boldsymbol{\lambda} = \{\lambda_j\}$, with $i, j = 1, \ldots, n+1$, the canonical basis for $\mathfrak{S}_{\mathbb{R}}$ and $\mathfrak{E}_{\mathbb{R}}$ as the same Euclidean space $\mathbb{R}^{n+1}$. The usual Bloch representation—in which the deterministic effect corresponds to the vector $\boldsymbol{\lambda}(e) = [0, \ldots, 0, 1] \in \mathbb{R}^{n+1}$—here becomes unsuitable. A more convenient representation of the simplex $\mathfrak{S}$ is the so called *probability simplex*, namely the n-dimensional polyhedra whose $(n+1)$ vertices correspond to the canonical basis vectors $\{\lambda_i\}$.[13] Naturally the cone of states $\mathfrak{S}_+$ becomes the $\mathbb{R}^{n+1}$

[13] Differently from the probabilistic models analysed until now, here the basis vectors $\{\lambda_i\} \in \mathfrak{S}_{\mathbb{R}}$ are true states of the classical theory.

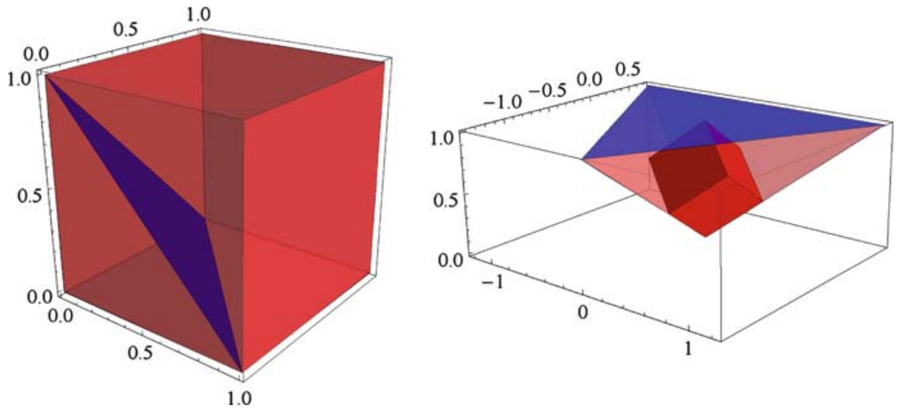


Fig. 7 *Left figure:* Probabilistic simplex representation of the trit system. The *triangle inside the cube* represents the simplex of states $\mathfrak{S}$ while the *transparent cube* is $\mathfrak{E}$. Both the cones $\mathfrak{S}_+$ and $\mathfrak{E}_+$ coincide with $\mathbb{R}^3_+$. *Right figure:* Bloch representation of the same trit system. The *transparent cone* represents both $\mathfrak{S}_+$ and $\mathfrak{E}_+$. The *triangle at the top* is the simplex of states $\mathfrak{S}$ while the convex of effects $\mathfrak{E}$ is the *inside solid cube*

positive orthant

$$\mathfrak{S}_+ = \mathbb{R}^{n+1}_+ = \left\{ \boldsymbol{l}(\omega) \in \mathbb{R}^{n+1} \mid \boldsymbol{l}(\omega) \succeq 0 \right\}, \quad \mathfrak{S} = \left\{ \boldsymbol{l}(\omega) \in \mathbb{R}^{n+1} \mid \boldsymbol{l}(\omega) \cdot \mathbf{1} = 1 \right\}, \tag{128}$$

where the symbol $\succeq$ denotes componentwise inequality[14] and $\mathbf{1}$ denotes the vector $[1, \dots, 1] \in \mathbb{R}^{n+1}$. In this representation the system is pointedly self-dual and the cone and set of effects are given by

$$\mathfrak{E}_+ = \mathbb{R}^{n+1}_+ = \left\{ \boldsymbol{\lambda}(a) \in \mathbb{R}^{n+1} \mid \boldsymbol{\lambda}(a) \succeq 0 \right\}, \quad \mathfrak{E} = \left\{ \boldsymbol{\lambda}(a) \in \mathbb{R}^{n+1} \mid 0 \preceq \boldsymbol{\lambda}(a) \preceq 1 \right\}. \tag{129}$$

The deterministic effect e, which must satisfy the condition $\omega(e) = 1 \ \forall \omega \in \mathfrak{S}$, and then $\lambda_i(e) = 1 \ \forall i$, is now represented by the vector $\boldsymbol{\lambda}(e) = \mathbf{1} \in \mathbb{R}^{n+1}$.

To clarify the situation we give a concrete representation of the classical theory with $\dim(\mathfrak{S}) = 2$. The simplex in dimension 2 is a triangle and the corresponding system is called **trit**, a generalization of the **bit** whose simplex of states is simply a segment. In the left Fig. 7 we show the probabilistic simplex representation of the trit system according to Eqs. (128) and (129). For completeness in Fig. 7 the usual Bloch representation of the same system is also reported.

[14] Componentwise or vector inequality in $\mathbb{R}^n$: $w \succeq v$ means $w_i \geqslant v_i$ for $i = 1, \dots, n$.

9.2 Simplex structure consequences

The first consequence of the simplex nature of $\mathfrak{S}$ is expressed in the following proposition.

Proposition 9.1 *A probabilistic theory has a simplex as local convex set of states if and only if the bipartite set of states is a simplex too.*

Proof Let $\mathfrak{S}$ be an n-dimensional simplex. We denote by $\omega_1, \omega_2, \ldots, \omega_{n+1}$ the vertices of $\mathfrak{S}$. Then the set of functionals $\{a_1, a_2, \ldots, a_n + 1\} \in \mathfrak{E}_{\mathbb{R}}$ such that

$$a_i(\omega_j) = \delta_{ij}, \tag{130}$$

are vertices of $\mathfrak{E}$. Notice that in the probability simplex representation the set of $\mathfrak{S}$-vertices $\mathrm{Extr}(\mathfrak{S}) = \{\omega_1, \omega_2, \ldots, \omega_{n+1}\}$ coincides with the orthonormal basis $\{\lambda_i\}$ for $\mathfrak{S}_{\mathbb{R}}$. The cone of transformations $\mathfrak{T}_+$ ($\dim(\mathfrak{T}_+) = (n+1)^2$) for a classical theory is the cone of positive maps, namely the maps preserving the local cone of states $\mathfrak{S}_+$. Then a map $\mathscr{A} \in \mathfrak{T}_+$ if and only if $\mathscr{A}\omega \in \mathfrak{S}_+ \ \forall \omega \in \mathrm{Extr}(\mathfrak{S})$, or, in the probabilistic simplex representation, $\mathscr{A}\lambda_i \in \mathfrak{S}_+ \ \forall \lambda_i$. Being $\{\lambda_i\}$ the canonical basis, it follows that $\mathfrak{T}_+$ includes all the transformations represented by a $(n+1) \times (n+1)$ matrix with all non negative elements. Then in the probabilistic simplex representation the extremal rays $\mathsf{Erays}(\mathfrak{T}_+)$ are generated by the $(n+1)^2$ matrices having an entry equal to one and all the other entries equal to zero. In a generic representation these rays are the transformations

$$\gamma\omega_i \otimes a_j \qquad \forall i, j = 1, \ldots, n+1, \ \forall \gamma \geqslant 0. \tag{131}$$

where γ is a multiplicative constant spanning the whole ray generated by the transformation $\omega_i \otimes a_j$. These maps send the convex set $\mathfrak{S}$ into an extremal ray of $\mathfrak{S}_+$. The preparationally faithful state of the theory Φ provides the isomorphisms $\mathfrak{E}_+ \simeq \mathfrak{S}_+$ ($\Phi(\cdot, a_j) = \omega_j$) and $\mathfrak{T}_+ \simeq \mathfrak{S}_+(\mathrm{AA})$. Remembering that a cone-isomorphism preserves the cone structure, from the $(n+1)^2$ extremal rays of $\mathfrak{T}_+$ in Eq. (131) we get the following $(n+1)^2$ extremal rays of $\mathfrak{S}_+(\mathrm{AA})$

$$\gamma\omega_i \otimes \omega_j \qquad \forall i, j = 1, \ldots, n+1, \gamma \geqslant 0. \tag{132}$$

Then the only bipartite pure states of the theory are the $(n+1)^2$ factorized states $\omega_i \otimes \omega_j$. In conclusion the bipartite set of states is a $((n+1)^2 - 1)$-dimensional convex set having $(n+1)^2$ vertices, namely a simplex.

The opposite implication—namely if $\mathfrak{S}(\mathrm{AA})$ is a simplex then $\mathfrak{S}$ is a simplex too—is trivial. Consider for example a $(n^2 - 1)$-dimensional bipartite simplex, then $\mathfrak{S}(\mathrm{AA})$ has only n^2 pure states. But $\mathfrak{S}$ cannot admit more than the n vertices generating the n^2 pure bipartite ones. Therefore $\mathfrak{S}$ is a simplex. □

This proposition has some interesting corollaries which show how, the classical theories are very special probabilistic theories.

Corollary 9.1 *The classical probabilistic theories are local.*

Proof A theory is said to be local if and only if it does not violate the CHSH inequality. The last proposition shows that if the local set of states is a simplex, then also the bipartite one is a simplex, and its vertices are factorized states. Then all the bipartite states are factorized probability rules which do not allow violations of the CHSH inequality. □

In the following corollary we give a property of the set of local automorphisms for a classical probabilistic theory. The group of automorphisms $\mathbf{G}_{\mathrm{A}}$ of a n-dimensional simplex is the *permutation group* $\mathbf{S}_{n+1}$, which contains the $(n+1)!$ different permutations of the set $\mathrm{Extr}(\mathfrak{S}) = \{\omega_1, \ldots, \omega_{n+1}\}$.

Corollary 9.2 *The local automorphisms of a classical probabilistic theory cannot be extremal transformations.*

Proof A general element of $\mathbf{G}_{\mathrm{A}} = \mathbf{S}_{n+1}$ can be identified by a set of indexes

$$J = \{j_1, \ldots, j_{n+1}\}, \tag{133}$$

representing a permutation of the set $\{1, \ldots, n+1\}$. The automorphism associated to such permutation is the map

$$\sum_{i=1,\ldots,n+1} \omega_{j_i} \otimes a_i, \tag{134}$$

which is manifestly a convex combination of the extremal transformations $\omega_{j_i} \otimes a_i$ given in Eq. (132) of Proposition 9.1.[15] □

Corollary 9.3 *The classical probabilistic theories do not satisfy Postulates PFAITH and PURIFY-1.*

Proof Notice that the identical transformation $\mathscr{I}$ is a particular permutation $\mathscr{I} \in \mathbf{S}_{n+1}$ and then a local automorphism of the classical theory. According to Corollary 9.2 the identical transformation cannot be atomic. On the other hand we know from Sect. 3.1 that Postulate PFAITH implies the atomicity of $\mathscr{I}$, whence it cannot be satisfied. For the same reason also Postulate PURIFY-1 does not hold. In fact, according to Lemma 5, it implies atomicity of the identical transformation. □

It is not surprising that Postulate PFAITH fails, since it assumes the existence of a pure preparationally faithful state. On the other hand, as showed in Proposition 9.1, the only pure bipartite states for a classical probabilistic theory are the factorized ones, which obviously do not achieve the isomorphism $\mathfrak{S}_+ \simeq \mathfrak{E}_+$, whence they are not preparationally faithful. Therefore, a preparationally faithful state cannot be pure and Postulate PFAITH fails. Also the impossibility of purifying a classical theory is

[15] In the probability simplex representation the automorphisms are the $(n+1) \times (n+1)$ permutation matrices. These are manifestly combinations of the extremal ones. In fact, as we have already shown, in the probability simplex representation the extremal transformations are the $(n+1) \times (n+1)$ matrices with an entry equal to one and the other entries equal to zero).

almost obvious, since there are not enough bipartite pure states to purify the continuous of internal points of the n-dimensional simplex $\mathfrak{S}$. More precisely, since the only bipartite pure states are the $(n+1)^2$ factorization of the $(n+1)$ pure states of $\mathfrak{S}$, no mixed state admits purification. A similar problem is suffered by the extended Popescu–Rohrlich model (see Sect. 5.6) where no mixed state in $\mathfrak{S}$, apart from its center χ, allows purification.

References

1. Birkhoff, G., von Neumann, J.: The logic of quantum mechanics. Ann. Math. **37**, 823 (1936)
2. Jordan, P., Neumann, J.von , Wigner, E.: On an algebraic generalization of the quantum mechanical formalism. Ann. Math. **35**, 29 (1934)
3. Varadarajan, V.S.: Probability in physics and a theorem on simultaneous observability. Comm. Pure Appl. Math. **15**, 189 (1962)
4. Mackey, G.W.: Mathematical Foundations of Quantum Mechanics. Benjamin, New York (1963)
5. Segal, I.E.: Postulates for general uantum mechanics. Ann. Math. **48**, 930 (1947)
6. Ludwig, G.: An axiomatic basis for quantum mechanics i: derivation of Hilbert space structure. Springer, Berlin (1985)
7. Hardy, L.: Quantum theory from five reasonable axioms, quant-ph/0101012v4 (2001)
8. Clifton, R., Bub, J., Halvorson, H.: Characterizing quantum theory in terms of information-theoretic constraints. Found. Phys. **33**, 1561 (2003)
9. D'Ariano, G.M.: Probabilistic theories: what is special about quantum mechanics. In: Bokulich, A., Jaeger, G. (eds.) Philosophy of Quantum Information and Entanglement. Cambridge University Press, Cambridge, UK, Also arXiv:0807.4383 (2010)
10. Popescu, S., Rohrlich, D.: Quantum non-locality as an axiom. Found. Phys. **24**, 379 (1994)
11. Wootters, W.K.: Why does nature likes the square root of negative one? Presentation at Perimeter, PIRSA:09110036 (2009)
12. Chiribella, G., D'Ariano, G.M., Perinotti, P.: Reversible realization of physical processes in probabilistic theories, arXiv:0908.1583
13. Abramsky, S., Coecke, B.: Proceedings of the 19th IEEE conference on Logic in Computer Science (LiCS'04), IEEE Computer Science Press (2004)
14. Chiribella, G., D'Ariano, G.M., Perinotti, P.: unpublished
15. Chiribella, G., D'Ariano, G.M., Perinotti, P.: Europhysics Lett. **83**, 30004 (2008)
16. Chiribella, G., D'Ariano, G.M., Perinotti, P.: Phys. Rev. A **80**, 022339 (2009)
17. Hardy, L.: J. Phys. A: Math. Theor. **40**, 3081 (2007)
18. Boyd, A., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge, UK (2004)
19. Barrett, J., Linden, N., Massar, S., Pironio, S., Popescu, S., Roberts, D.: Non-local correlations as an information theoretic resource. Phys. Rev. A **71**, 22101 (2005)
20. Cirel'son, B.S.: Quantum generalizations of bell's inequality. Lett. Math. Phys. **4**, 93 (1980)