A derivation of quantum theory from physical requirements

Lluís Masanes
ICFO-Institut de Ciencies Fotoniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain

Markus P. Müller
Institute of Mathematics, Technical University of Berlin, 10623 Berlin, Germany, and
Institute of Physics and Astronomy, University of Potsdam, 14476 Potsdam, Germany

(Dated: October 5, 2010)

Quantum theory is derived from five requirements which are imposed on the framework of generalized probabilistic theories. These requirements are simple and have a clear physical meaning, in terms of basic operational procedures. They do not refer to the mathematical structure and representation of states and measurements.

I. INTRODUCTION

Quantum theory is usually formulated by postulating the mathematical structure and representation of states, measurements, and transformations. The general physical consequences that follow (possibility of local tomography, violation of Bell-type inequalities [1], factorization of integers in polynomial time [2], etc.) come as theorems which use the postulates as premises. In this work this procedure is reversed: we impose five simple physical requirements, and this suffices to single out quantum theory and derive its mathematical formalism uniquely. This is more similar to the usual formulation of Special Relativity, where two simple physical requirements are used to derive the mathematical structure of Minkowski space-time and its transformations.

The requirements can be schematically stated as:

1. In systems that carry one bit of information, each state is characterized by a finite set of outcome probabilities.

2. The state of a composite system is characterized by the statistics of measurements on the individual components.

3. All systems that effectively carry the same amount of information have equivalent state spaces.

4. Any pure state of a system can be reversibly transformed into any other.

5. In systems that carry one bit of information, all mathematically well-defined measurements are allowed by the theory.

These requirements are imposed on the framework of generalized probabilistic theories [3–9], which already assumes that some operational notions (preparation, mixture, measurement, and counting relative frequencies of measurement outcomes) make sense. Due to its conceptual simplicity, this framework leaves room for an infinitude of possible theories, allowing for weaker- or stronger-than-quantum non-locality [6, 10–14]. In this work, we show that quantum theory (QT) and classical probability theory (CPT) are very special among those theories: they are the only general probabilistic theories that satisfy the five requirements stated above. In addition, we show below that this claim may be reformulated in a way which makes Requirement 5 unnecessary.

The non-uniqueness of the solution is not a problem, since CPT is embedded in QT. One can also proceed as Hardy in [4]: if Requirement 4 is strengthened by imposing continuity of the reversible transformations, then CPT is ruled out and QT is the only theory satisfying the requirements. This strengthening can be justified by the continuity of time evolution in physical systems.

It is conceivable that in the future, another theory may replace or generalize QT. Such a theory must violate at least one of our assumptions. The clear meaning of our requirements allows to straightforwardly explore potential features of such a theory. The relaxation of each of the requirements constitutes a different way to go beyond QT. Most attempts to modify QT have been based on altering its mathematical formalism [15]. A derivation of QT in terms of physical requirements may provide a more transparent approach for this endeavor.

The search for alternative axiomatizations of QT is an old topic, which has been approached in many different ways: extending propositional logic [7, 8], using operational primitives [3–6, 9], searching for information-theoretic principles [5, 6, 10, 11, 16–18], building upon the phenomenon of quantum nonlocality [6, 10–13]. Alfsen and Shultz [19] have accomplished a complete characterization of the state spaces of QT from a geometric point of view, but the result does not seem to have an immediate physical meaning. In particular, the fact that the state space of a generalized bit is a three-dimensional ball is an assumption there, while here it is derived from physical requirements.

This work is particularly close to [4, 16], from where it takes some material. More concretely, the multiplicativity of capacities and the Simplicity Axiom from [4] are replaced by Requirement 5. In comparison with [16], the fact that each state of a generalized bit is the mixture of two distinguishable ones, the group of reversible transformations and its orthogonality, and the multiplicativity of capacities, are replaced by Requirement 5.

Summary of the paper. Section II contains an introduc-

tion to the framework of generalized probabilistic theories, where some elementary results are stated without proof. In Section III the five requirements and their significance are explained in full detail, and it is argued that Requirement 5 is in a sense unnecessary. Section IV is the core of this work. It contains the characterization of all theories compatible with the requirements, concluding that the only possibilities are CPT and QT. The Conclusion (Section V) recapitulates the results and adds some remarks. The Appendix contains all lemmas and their proofs.

II. GENERALIZED PROBABILISTIC THEORIES

In CPT there can always be a joint probability distribution for all random variables under consideration. The framework of generalized probabilistic theories (GPTs), also called convex operational framework, generalizes this by allowing the possibility of random variables that cannot have a joint probability distribution, or cannot be simultaneously measured (like noncommuting observables in QT).

This framework assumes that at some level there is a classical reality, where it makes sense to talk about experimentalists performing basic operations such as: preparations, mixtures, measurements, and counting relative frequencies of outcomes. These are the primary concepts of this framework. It also provides a unified way (within all theories assuming that classical level) to mathematically represent states, transformations and measurements. A particular GPT specifies which of these are allowed, but it does not tell their correspondence with actual experimental setups. On its own, a GPT can still make nontrivial predictions like: the maximal violation of a Bell inequality [1], the complexity-theoretic computational power [2, 15], and in general, all information-theoretic properties of the theory [6].

The framework of GPTs can be stated in different ways, but all lead to the same formalism [3–9]. This formalism is presented in this section at a very basic level, providing some elementary results without proofs.

A. States

Definition of system. To a setup like FIG. 1 we associate a system if for each configuration of the preparation, transformation and measurement devices, the relative frequencies of the outcomes tend to a unique probability distribution (in the large sample limit).

The probability of a measurement outcome x is denoted by $p(x)$. This outcome can be associated to a binary measurement which tells whether x happens or not (this second event $\bar{x}$ has probability $p(\bar{x}) = 1 - p(x)$). The above definition of system allows to associate to each preparation procedure a list of probabilities for the outcomes of all measurement that can be performed on a system. As we show in Subsection IV C below, our requirements imply that all these probabilities $p(x)$ are determined by a finite set of them; the smallest such set is used to represent the state

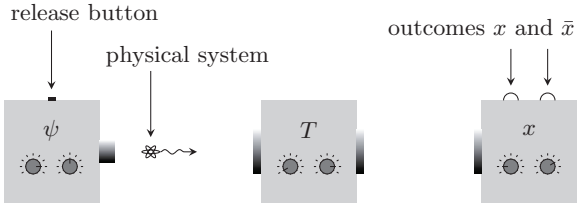

FIG. 1: This is a pictorial representation of a general experimental set up; with preparation, transformation and measurement devices (from left to right). As soon as the release button is pressed, the preparation device outputs a physical system in the state specified by the knobs. The next device performs the transformation specified by its knobs (which in particular can be "do nothing"). The device on the right performs the measurement specified by its knobs, and the outcome (x or $\bar{x}$) is indicated by the corresponding light.

$$\psi = \begin{bmatrix} 1 \\ p(x_1) \\ \vdots \\ p(x_d) \end{bmatrix} = \begin{bmatrix} \psi^0 \\ \psi^1 \\ \vdots \\ \psi^d \end{bmatrix} \in \mathcal{S} \subset \mathbb{R}^{d+1}. \qquad (1)$$

The measurements that characterize the state $x_1, \ldots, x_d$ are called *fiducial*, and in general, there is more than one set of them (for example, a $\frac{1}{2}$-spin particle in QT is characterized by the spin in any 3 linearly-independent directions). The redundant component $\psi^0 = 1$ is reminiscent of QT, where one of the diagonal entries of a density matrix is redundant, since they sum up to 1. In fact $\psi^0 \neq 1$ is sometimes used to represent unnormalized states, but not in this paper, where only normalized states are considered. The redundant component ψ^0 allows to use the tensor-product formalism in composite systems (Subsection II D), which simplifies the notation.

The set of all allowed states $\mathcal{S}$ is convex [20], because if $\psi_1, \psi_2 \in \mathcal{S}$ then one can prepare ψ_1 with probability q and ψ_2 with probability $1 - q$, effectively preparing the state $q\psi_1 + (1-q)\psi_2$. The number of fiducial probabilities d is equal to the (affine) dimension of $\mathcal{S}$, otherwise one fiducial probability would be functionally related to the others, and hence redundant.

Suppose there is a $\mathbb{R}^{d+1}$-vector $\psi \notin \mathcal{S}$ which is in the topological closure of $\mathcal{S}$ – that is, ψ can be approximated by states $\psi' \in \mathcal{S}$ to arbitrary accuracy. Since there is no observable physical difference between *perfect preparation* and *arbitrarily good preparation*, we will consider ψ to be a valid state and add it to the state space. This does not change the physical predictions of the theory, but it has the mathematical consequence that state spaces become topologically closed. Since state vectors (1) are bounded, and we are in finite dimensions (shown in Sub-

section IV C), state spaces $\mathcal{S}$ are *compact* convex sets [20].

The *pure states* of a state space $\mathcal{S}$ are the ones that cannot be written as mixtures: $\psi \neq q\psi_1 + (1-q)\psi_2$ with $\psi_1 \neq \psi_2$ and $0 < q < 1$. Since $\mathcal{S}$ is compact and convex, all states are mixtures of pure states [20].

B. Measurements

The probability of measurement outcome x when the system is in state $\psi \in \mathcal{S}$ is given by a function $\Omega_x(\psi)$. Suppose the system is prepared in the mixture $q\psi_1 + (1-q)\psi_2$, then the relative frequency of outcome x does not depend on whether the label of the actual preparation ψ_k is ignored before or after the measurement, hence

$$\Omega_x\big(q\psi_1 + (1-q)\psi_2\big) = q\,\Omega_x(\psi_1) + (1-q)\,\Omega_x(\psi_2)\ .$$

This means that the function Ω_x is affine on $\mathcal{S}$. The redundant component ψ^0 in (1) allows to write this function as a linear map $\Omega_x : \mathbb{R}^{d+1} \to \mathbb{R}$ [3, 6].

An *effect* is a linear map $\Omega : \mathbb{R}^{d+1} \to \mathbb{R}$ such that $\Omega(\psi) \in [0,1]$ for all states $\psi \in \mathcal{S}$. Every function Ω_x associated to an outcome probability $p(x)$ is an effect. The converse is not necessarily true: the framework of GPTs allows to construct theories where some effects do not represent any of the measurements allowed by the theory. These restrictions are analogous to superselection rules, where some (mathematically well-defined) states are not allowed by the theory. This is related to Requirement 5. A *tight effect* Ω is one for which there are two states $\psi_0, \psi_1 \in \mathcal{S}$ satisfying $\Omega(\psi_0) = 0$ and $\Omega(\psi_1) = 1$.

An n-outcome measurement is specified by n effects $\Omega_1, \ldots, \Omega_n$ such that $\Omega_1(\psi) + \cdots + \Omega_n(\psi) = 1$ for all $\psi \in \mathcal{S}$. The number $\Omega_a(\psi)$ is the probability of outcome a when the measurement is performed on the state ψ. The states $\psi_1, \ldots, \psi_n$ are *distinguishable* if there is an n-outcome measurement such that $\Omega_a(\psi_b) = \delta_{a,b}$, where $\delta_{a,b} = 1$ if $a = b$, and $\delta_{a,b} = 0$ if $a \neq b$.

The *capacity* of a state space $\mathcal{S}$ is the size of the largest family of distinguishable states, and is denoted by c. This is the amount of classical information that can be transmitted by the corresponding type of system, in a single-shot error-free procedure. (In QT the capacity of a system is the dimension of its corresponding Hilbert space; which must not be confused with the dimension of the state space $d = c^2 - 1$, that is, the set of positive unit-trace complex matrices.) A *complete measurement* on $\mathcal{S}$ is one capable of distinguishing c states.

C. Transformations

Each type of system has associated to it: a state space, a set of measurements, and a set of transformations. A transformation T is a map $T : \mathcal{S} \to \mathcal{S}$. Similarly as for measurements, if a state is prepared as a mixture $q\psi_1 + (1-q)\psi_2$, it does not matter whether the label of the actual preparation ψ_k is ignored before or after the transformation. Hence

$$T\left(q\psi_1 + (1-q)\psi_2\right) = qT(\psi_1) + (1-q)T(\psi_2)\ ,$$

which implies that T is an affine map. The redundant component ψ^0 in (1) allows to extend T to a linear map $T : \mathbb{R}^{d+1} \to \mathbb{R}^{d+1}$ [3, 6].

A transformation T is reversible if its inverse T^{-1} exists and belongs to the set of transformations allowed by the theory. The set of (allowed) reversible transformations of a particular state space $\mathcal{S}$ forms a group $\mathcal{G}$. For the same reason as for the state space itself, we will assume that the group of reversible transformations is topologically closed. Previously we have seen that a state space $\mathcal{S}$ is bounded, hence the corresponding group of transformations $\mathcal{G}$ is bounded, too. In summary, groups of transformations are compact [21].

D. Composite systems

Definition of composite system. Two systems A, B constitute a composite system, denoted AB, if a measurement for A together with a measurement for B uniquely specifies a measurement for AB. This means that if x and y are measurement outcomes on A and B respectively, the pair (x, y) specifies a unique measurement outcome on AB, whose probability distribution $p(x, y)$ does not depend on the temporal order in which the subsystems are measured.

The fact that subsystems are themselves systems implies that each has a well-defined reduced state ψ_A, ψ_B which does not depend on which transformations and measurements are performed on the other subsystem (see definition of system in Subsection II A). This is often referred to as no-signaling. Let $x_1, \ldots, x_{d_A}$ be the fiducial measurements of system A, and $y_1, \ldots, y_{d_B}$ the ones of B. The no-signaling constraints are

$$\begin{array}{rcl} p(x_i) & = & p(x_i, y_j) + p(x_i, \bar{y}_j) \\ p(y_i) & = & p(x_i, y_j) + p(\bar{x}_i, y_j) \end{array} \qquad (2)$$

for all i, j.

An assumption which is often postulated additionally in the GPT context is Requirement 2, which says that the state of a composite system is completely characterized by the statistics of measurements on the subsystems, that is, $p(x, y)$. This and no-signaling (2) imply that states in AB can be represented on the tensor product vector

space [3] as

$$\psi_{AB} = \begin{pmatrix} 1 \\ \vdots \\ p(x_i) \\ \vdots \\ p(y_j) \\ \vdots \\ p(x_i, y_j) \\ \vdots \end{pmatrix} \in \mathcal{S}_{AB} \subset \mathbb{R}^{d_A+1} \otimes \mathbb{R}^{d_B+1}. \quad (3)$$

The joint probability of two arbitrary local measurement outcomes x, y is given by

$$p(x, y) = (\Omega_x \otimes \Omega_y)(\psi_{AB}) \ , \quad (4)$$

where Ω_x is the effect representing x in A, that is $p(x) = \Omega_x(\psi_A)$, and analogously for Ω_y [3]. (The term "local" is used when referring to subsystems, and has nothing to do with spatial locations.) In other words, if $\{\Omega_1^A, \ldots, \Omega_n^A\}$ is an n-outcome measurement on A, and $\{\Omega_1^B, \ldots, \Omega_m^B\}$ is an m-outcome measurement on B, then $\{\Omega_a^A \otimes \Omega_b^B |\ a = 1, \ldots, n\,; b = 1, \ldots, m\}$ defines a measurement on AB with nm outcomes. Local transformations act on the global state as

$$\psi_{AB} \ \rightarrow \ (T_A \otimes T_B)(\psi_{AB}) \ , \quad (5)$$

where T_A is the matrix that represents the transformation in A, and analogously for T_B [3]. The reduced states

$$\psi_A = \begin{pmatrix} 1 \\ \vdots \\ p(x_i) \\ \vdots \end{pmatrix} \ , \quad \psi_B = \begin{pmatrix} 1 \\ \vdots \\ p(y_j) \\ \vdots \end{pmatrix} \ , \quad (6)$$

are obtained from ψ_{AB} by picking the right components (3). Alternatively, reduced states can be defined by $\Omega_A(\psi_A) = (\Omega_A \otimes \mathbf{1})(\psi_{AB})$ for any effect Ω_A in A, where $\mathbf{1}(\psi_B) = \psi_B^0$ is the unit effect. The reduced state ψ_A must belong to the state space of subsystem A, denoted $\mathcal{S}_A$, and any state in $\mathcal{S}_A$ must be the reduction of a state from $\mathcal{S}_{AB}$. (Analogously for subsystem B.) This implies that all product states

$$\psi_{AB} = \psi_A \otimes \psi_B \quad (7)$$

are contained in $\mathcal{S}_{AB}$ [3], and similarly, all tensor products of local measurements and transformations are allowed on AB.

Given two fixed state spaces $\mathcal{S}_A$ and $\mathcal{S}_B$, the previous discussion imposes constraints on the state space of the composite system $\mathcal{S}_{AB}$. However, there are still many different possibilities for how to define $\mathcal{S}_{AB}$.

Nothing prevents Bob's system from being composite itself; hence one can recursively extend the definition of composite system and formulas (3), (4), (5), and (7) to more parties.

E. Equivalent state spaces

Let $\mathcal{L} : \mathcal{S} \rightarrow \mathcal{S}'$ be an invertible affine map. If all states are transformed as $\psi \rightarrow \mathcal{L}(\psi)$, and all effects on $\mathcal{S}$ are transformed as $\Omega \rightarrow \Omega \circ \mathcal{L}^{-1}$, then the outcome probabilities $\Omega(\psi)$ are kept unchanged. Analogously, if all transformations on $\mathcal{S}$ are mapped as $T \rightarrow \mathcal{L} \circ T \circ \mathcal{L}^{-1}$ then their action on the states is the same. The new state space $\mathcal{S}'$, together with the transformed effects and transformations, is then just a different representation of $\mathcal{S}$. In this case, we call $\mathcal{S}$ and $\mathcal{S}'$ *equivalent*. In the new representation, the entries of ψ need not be probabilities as in (1), but it may have other advantages. In this work, several representations are used.

In the standard formalism of QT, states are represented by density matrices, however they can also be represented as in (1).

Changing the set of fiducial measurements is a particular type of $\mathcal{L}$-transformation. For example, if the components of the Bloch vector (of a quantum spin-$\frac{1}{2}$ particle) correspond to spin measurements in non-orthogonal directions, then the Bloch sphere becomes an ellipsoid.

F. Instances of generalized probabilistic theories

QT is an instance of GPT, and can be specified as follows. The state space $\mathcal{S}_c$ with capacity c is equivalent to the set of complex $c \times c$-matrices ρ such that $\rho \geq 0$ and $\mathrm{tr}\rho = 1$. This set has dimension $d_c = c^2 - 1$, and its pure states are constrained to be rank-one. The effects on $\mathcal{S}_c$ have the form $\Omega(\rho) = \mathrm{tr}(M\rho)$, where M is a complex $c \times c$-matrix such that $0 \leq M \leq \mathbb{I}$. The reversible transformations act as $\rho \rightarrow V\rho V^\dagger$ with $V \in \mathrm{SU}(c)$. The capacity of a composite system AB is the product of the capacities for the subsystems $c_{AB} = c_A c_B$.

CPT is another instance of GPT, and can be specified as follows. The state space $\mathcal{S}_c$ with capacity c is equivalent to the set of c-outcome probability distributions $[p(1), \ldots, p(c)]$, which has dimension $d_c = c - 1$ (in geometric terms, each $\mathcal{S}_c$ is a simplex). The pure states are the deterministic distributions $p(a) = \delta_{a,b}$ with $b = 1, \ldots, c$. The c-outcome measurement with effects $\Omega_a(\psi) = p(a)$ for $a = 1, \ldots, c$, distinguishes the c pure states, hence it is complete. Any other measurement is a function of this one. The reversible transformations act by permuting the entries of the state $[p(1), \ldots, p(c)]$. The capacity of a composite system is also $c_{AB} = c_A c_B$. Note that CPT can be obtained by restricting the states of QT to diagonal matrices. In other words, CPT is embedded in QT.

An instance of GPT that is not observed in nature is *generalized no-signaling theory* [6], colloquially called *boxworld*. By definition, state spaces contain all correlations (3) satisfying the no-signaling constraints (2). Such state spaces have finitely many pure states, and some of them violate Bell inequalities stronger than any quantum state [12]. The effects in boxworld are all generated by

products of local effects. The group of reversible transformations consists only of relabellings of local measurements and their outcomes, permutations of subsystems, and combinations thereof [14].

III. THE REQUIREMENTS

This section contains the precise statement of the requirements, each followed by explanations about its significance. At the end, the claim of this work is reformulated in a way that makes Requirement 5 unnecessary.

Requirement 1 (Finiteness). *A state space with capacity $c = 2$ has finite dimension d.*

If this did not hold, the characterization of a state of a generalized bit would require infinitely many outcome probabilities. Unless additional constraints (like energy bounds) were imposed, state estimation would become impossible. It is shown below that this requirement, together with the others, implies that all state spaces with finite capacity c have finite dimension.

Requirement 2 (Local tomography). *The state of a composite system AB is completely characterized by the statistics of measurements on the subsystems A, B.*

In other words, state tomography [3] can be performed locally. This is equivalent to the constraint

$$(d_{AB} + 1) = (d_A + 1)(d_B + 1) \tag{8}$$

[3, 4]. This requirement can be recursively extended to more parties by letting subsystems A, B to be themselves composite.

Requirement 3 (Equivalence of subspaces). *Let $\mathcal{S}_c$ and $\mathcal{S}_{c-1}$ be systems with capacities c and $c-1$, respectively. If $\Omega_1, \ldots, \Omega_c$ is a complete measurement on $\mathcal{S}_c$, then the set of states $\psi \in \mathcal{S}_c$ with $\Omega_c(\psi) = 0$ is equivalent to $\mathcal{S}_{c-1}$.*

The notions of complete measurements and equivalent state spaces are defined in Subsections II B and II E. In particular, equivalence of $\mathcal{S}_{c-1}$ and

$$\mathcal{S}'_{c-1} := \{\psi \in \mathcal{S}_c : \Omega_c(\psi) = 0\} \subset \mathcal{S}_c \tag{9}$$

implies that all measurements and reversible transformations on one of them can be implemented on the other.

This requirement, first introduced in [4], implies that all state spaces with the same capacity are equivalent: if $\mathcal{S}_{c-1}$ and $\tilde{\mathcal{S}}_{c-1}$ are state spaces with capacity $c-1$, then both are equivalent to (9), hence they are equivalent to each other. In other words, the only property that characterizes the type of system is the capacity for carrying information. If we start with $\mathcal{S}_c$ and apply Requirement 3 recursively, we get a more general formulation: consider any subset of outcomes $\{a_1, \ldots, a_{c'}\} \subseteq \{1, \ldots, c\}$ of the complete measurement $\Omega_1, \ldots, \Omega_c$, then the set of states $\psi \in \mathcal{S}_c$ with

$$\Omega_{a_1}(\psi) + \cdots + \Omega_{a_{c'}}(\psi) = 1 \tag{10}$$

is equivalent to the state space $\mathcal{S}_{c'}$ with capacity c'. This provides an onion-like structure for all state spaces $\mathcal{S}_1 \subset \mathcal{S}_2 \subset \mathcal{S}_3 \subset \cdots$

Requirement 3 is related to the notion of information. It allows to refer to states abstractly, without specifying the type of system they represent. It allows the existence of a unit of information (for instance, a system with $c = 2$) such that a sufficient number of copies of it can store the state of any other system, which later can be retrieved with in principle no losses. It allows for the existence of efficient universal simulators within the theory.

Requirement 4 (Symmetry). *For every pair of pure states $\psi_1, \psi_2 \in \mathcal{S}$ there is a reversible transformation G mapping one onto the other: $G(\psi_1) = \psi_2$.*

The set of reversible transformations of a state space $\mathcal{S}_c$ forms a group, denoted $\mathcal{G}_c$. This group endows $\mathcal{S}_c$ with a symmetry, which makes all pure states equivalent. A group $\mathcal{G}_c$ is said to be *continuous* if it is topologically connected: any transformation is the composition of many infinitesimal ones [21]. Hardy invokes the continuity of time-evolution in physical systems to justify the continuity of reversible transformations [3, 4]; in this case, state spaces $\mathcal{S}_c$ must have infinitely-many pure states; this rules out CPT and singles out QT. However, all the analysis in this work is done without imposing continuity, since we find it very interesting that the only theory with state spaces having finitely-many pure states, and satisfying the requirements, is CPT.

Requirement 5 (All measurements allowed). *All tight effects on $\mathcal{S}_2$ correspond to allowed measurements.*

It is shown below that, in combination with the other requirements, this implies that all effects on all state spaces (with arbitrary c) can be measured.

Consider a theory according to which some effects on $\mathcal{S}_2$ are fundamentally impossible to measure. Define an idealized version of this theory, which allows all such (impossible) measurements. If this idealized theory satisfies Requirements 1, 2, 3, 4 then it must be QT or CPT, and the original theory must be a restriction of QT or CPT [22]. Then, Requirement 5 is unnecessary if the claim of this work is formulated as:

Any theory which has an idealized version of itself satisfying Requirements 1, 2, 3 and 4 is embedded in QT or CPT.

IV. CHARACTERIZATION OF ALL THEORIES SATISFYING THE REQUIREMENTS

A. The maximally-mixed state

We use the following notation: the system with capacity c has state space $\mathcal{S}_c$ with dimension d_c and group of reversible transformations $\mathcal{G}_c$. The group $\mathcal{G}_c$ is compact (Section II C), and hence, has a normalized invariant

Haar measure [23]. This allows to define the maximally-mixed state

$$\mu_c = \int_{\mathcal{G}_c} G(\psi)\, dG \ \in \mathcal{S}_c \ , \tag{11}$$

where $\psi \in \mathcal{S}_c$ is an arbitrary pure state. It follows from Requirement 4 that the resulting state μ_c does not depend on the choice of the pure state ψ. By construction, the maximally-mixed state is invariant:

$$G(\mu_c) = \mu_c \ \text{ for all } \ G \in \mathcal{G}_c \ . \tag{12}$$

Moreover, Lemma 1 shows that it is the only invariant state in $\mathcal{S}_c$.

B. The generalized bit

A generalized bit is a system with capacity two. For any state $\psi \in \mathcal{S}_2$ in the standard representation (1), its Bloch representation is defined by

$$\hat{\psi} = 2 \begin{bmatrix} p(x_1) - \mu_2^1 \\ \vdots \\ p(x_{d_2}) - \mu_2^{d_2} \end{bmatrix} \in \hat{\mathcal{S}}_2 \ \subset \mathbb{R}^{d_2} \ . \tag{13}$$

States in the Bloch representation do not have the redundant component ψ^0, so equations (4, 5, 7) become less simple. The invertible map $\mathcal{L} : \mathcal{S}_2 \to \hat{\mathcal{S}}_2$ is affine but not linear; hence, effects Ω in the Bloch representation ($\hat{\Omega} = \Omega \circ \mathcal{L}^{-1}$) are affine but not necessarily linear. The same applies to transformations ($\hat{G} = \mathcal{L} \circ G \circ \mathcal{L}^{-1}$), however, the maximally-mixed state in the Bloch representation is the null vector $\hat{\mu}_2 = \mathbf{0}$, therefore (12) becomes $\hat{G}(\mathbf{0}) = \mathbf{0}$, which implies that $\hat{G}$ acts linearly (as a matrix).

Theorem 1. *A state in $\hat{\mathcal{S}}_2$ is pure if and only if it belongs to the boundary $\partial\hat{\mathcal{S}}_2$.*

Proof. In any convex set, pure states belong to the boundary [20]. Let us see the converse.

It is shown in [24] that any compact convex set has a supporting hyperplane containing exactly one point of the set. Translated to our language: there is a tight effect $\hat{\Omega}_{\rm one}$ on $\hat{\mathcal{S}}_2$ such that only one state $\hat{\varphi}_{\rm one} \in \hat{\mathcal{S}}_2$ satisfies $\hat{\Omega}_{\rm one}(\hat{\varphi}_{\rm one}) = 1$; this is illustrated in FIG. 2. According to Requirement 5, the effect $\hat{\Omega}_{\rm one}$ corresponds to a valid measurement outcome, and so does $\hat{\mathbf{1}} - \hat{\Omega}_{\rm one}$, where $\hat{\mathbf{1}}(\hat{\psi}) = 1$ for all $\hat{\psi} \in \hat{\mathcal{S}}_2$. Thus, the two effects $\hat{\Omega}_{\rm one}$ and $\hat{\mathbf{1}} - \hat{\Omega}_{\rm one}$ define a complete measurement on $\hat{\mathcal{S}}_2$. Imposing Requirement 3 on the single outcome $\hat{\Omega}_{\rm one}$ constrains the state space with unit capacity $\hat{\mathcal{S}}_1$ to contain only one state.

Suppose there is a point in the boundary $\hat{\varphi}_{\rm mix} \in \partial\hat{\mathcal{S}}_2$ which is not pure: $\hat{\varphi}_{\rm mix} = q\hat{\varphi}_1 + (1-q)\hat{\varphi}_2$ with $\hat{\varphi}_1 \neq \hat{\varphi}_2$ and $0 < q < 1$. Every point in the boundary of a compact convex set has a supporting hyperplane which contains it [20]. In our language: there is a tight effect $\hat{\Omega}$ on $\hat{\mathcal{S}}_2$ such that $\hat{\Omega}(\hat{\varphi}_{\rm mix}) = 1$. The affine function $\hat{\Omega}$ is bounded: $\hat{\Omega}(\hat{\varphi}) \leq 1$ for any $\hat{\varphi} \in \hat{\mathcal{S}}_2$, which implies $\hat{\Omega}(\hat{\varphi}_1) = \hat{\Omega}(\hat{\varphi}_2) = 1$; this is illustrated in FIG. 2. Like $\hat{\Omega}_{\rm one}$, the effect $\hat{\Omega}$ defines a complete measurement, and Requirement 3 can be imposed on the single outcome $\hat{\Omega}$, implying that $\hat{\mathcal{S}}_1$ contains more than one state. This is in contradiction with the previous paragraph; hence, all points in the boundary are pure. □

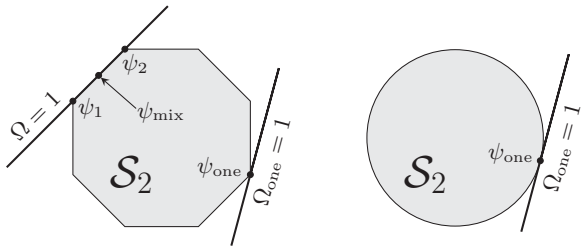

FIG. 2: The left figure is a state space whose boundary consists of facets (like $\Omega = 1$). Each facet contains infinitely many states ($\Omega = 1$ contains ψ_1, ψ_2 and all $\psi_{\rm mix} = q\psi_1 + (1-q)\psi_2$). The right figure is a state space whose boundary has no facets. Any state space has supporting hyperplanes containing a unique state (like $\Omega_{\rm one} = 1$ in both figures).

For the case $d_2 = 1$, the state space $\mathcal{S}_2$ is a segment (a 1-dimensional ball), hence the previous and next theorems are trivial. For $d_2 > 1$, the previous theorem implies that $\mathcal{S}_2$ contains infinitely-many pure states. The next theorem recovers the (quantum-like) Bloch sphere with a yet unknown dimension d_2.

Theorem 2. *There is a set of fiducial measurements for which $\hat{\mathcal{S}}_2$ is a d_2-dimensional unit ball.*

Proof. Lemma 2 shows that there is an invertible real matrix S such that for each $\hat{G} \in \hat{\mathcal{G}}_2$ the matrix $S\hat{G}S^{-1}$ is orthogonal. Let us redefine the set $\hat{\mathcal{S}}_2$ by transforming the states as $\hat{\varphi} \to \hat{\varphi}' = qS\hat{\varphi}$, where the number $q > 0$ is chosen such that all pure states are unit vectors $|\hat{\varphi}'|^2 = \hat{\varphi}'^{\rm T}\hat{\varphi}' = 1$. This is possible because in the transformed state space, all pure states are related by orthogonal matrices (SGS^{-1}) which preserve the norm. Since Theorem 1 also applies to the redefined set $\hat{\mathcal{S}}_2'$, it must be a unit ball. In what follows we define a new set of fiducial measurements x_i' such that the Bloch representation (13) associated to the new fiducial probabilities $p(x_i')$ coincides with the redefinition $\hat{\varphi}'$.

Requirement 5 tells that in $\hat{\mathcal{S}}_2'$, all tight effects are allowed measurements. For each unit vector $\hat{\nu} \in \mathbb{R}^{d_2}$ the function $\hat{\Omega}_{\hat{\nu}}(\hat{\varphi}') = (1 + \hat{\nu}^{\rm T}\hat{\varphi}')/2$ is a tight effect on the unit ball, and conversely, all tight effects on the unit ball are of this form. The new set of fiducial measurements

x'_i has effects $\hat{\Omega}_{x'_i} = \hat{\Omega}_{\hat{\nu}_i}$, where

$$\hat{\nu}_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \ \hat{\nu}_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \ \dots, \ \hat{\nu}_{d_2} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \tag{14}$$

is a fixed orthonormal basis for $\mathbb{R}^{d_2}$. For any state $\hat{\varphi}'$ the new fiducial probabilities are $p(x'_i) = \Omega_{x'_i}(\hat{\varphi}') = (1 + \hat{\varphi}'^i)/2$, which implies $\hat{\varphi}'^i = 2[p(x'_i) - 1/2]$. This is just (13) with the new fiducial measurements (note that $\hat{\mu}'_2 = \mathbf{0}$ and $\mu'^i_2 = \hat{\Omega}_{x'_i}(\mathbf{0}) = 1/2$). □

In the rest of the paper, we will use the representation derived in Theorem 2 above, where the generalized bit is represented by a unit ball. Moreover, we will drop the prime in $\hat{\mathcal{S}}'_2$, x'_i, $\hat{\varphi}'$ used in the proof, and simply write $\hat{\mathcal{S}}_2$, x_i, $\hat{\varphi}$.

As argued above, for each pure state $\varphi \in \mathcal{S}_2$ there is a binary measurement with associated effect

$$\Omega_\varphi(\psi) = (1 + \hat{\varphi}^{\mathrm{T}}\hat{\psi})/2 \ , \tag{15}$$

such that $\hat{\Omega}_\varphi(\hat{\varphi}) = 1$ and $\hat{\Omega}_\varphi(-\hat{\varphi}) = 0$. In summary, there is a correspondence between tight effects and pure states in $\mathcal{S}_2$, and each pure state belongs to a distinguishable pair $\{\hat{\varphi}, -\hat{\varphi}\}$.

C. Capacity and dimension

Requirements 1, 2 and 3 imply that a state space with finite capacity c has finite dimension d_c, which generalizes Requirement 1. To see this, consider a system composed of m generalized bits, with state space denoted by $\mathcal{S}_{2\times m}$. Since d_2 is finite, equation (8) implies that $\mathcal{S}_{2\times m}$ has finite dimension. Thus, its capacity, denoted c_m, must be finite. Since systems with the same capacity are equivalent, we must have $c_m \neq c_n$ for $m \neq n$, and the sequence of integers $c_1, c_2, \dots$ is unbounded. For any capacity c there is a value of m such that $c \leq c_m$, hence by Requirement 3 we have $\mathcal{S}_c \subset \mathcal{S}_{2\times m}$, which implies that $\mathcal{S}_c$ is finite-dimensional.

In QT, the maximally-mixed state (11) has two convenient properties. First property: if μ_A and μ_B are the maximally-mixed states of systems A and B, then the maximally-mixed state of the composite system AB is

$$\mu_{AB} = \mu_A \otimes \mu_B \ . \tag{16}$$

Second property: in the state space $\mathcal{S}_c$, there are c pure distinguishable states $\psi_1, \dots, \psi_c \in \mathcal{S}_c$ such that

$$\mu_c = \frac{1}{c}\sum_{a=1}^{c} \psi_a \ . \tag{17}$$

Lemmas 3 and 5 show that these two properties hold for every theory satisfying our requirements. The following theorem exploits these properties to show that the capacity is multiplicative (one of the axioms in [4]).

Theorem 3. *If c_A and c_B are the capacities of systems A and B, then the capacity of the composite system AB is*

$$c_{AB} = c_A c_B \ . \tag{18}$$

Proof. Equation (17) allows to write the maximally-mixed states of systems A and B as

$$\mu_A = \frac{1}{c_A}\sum_{a=1}^{c_A} \varphi^A_a \ , \quad \mu_B = \frac{1}{c_B}\sum_{b=1}^{c_B} \varphi^B_b \ ,$$

where $\varphi^A_1, \dots, \varphi^A_{c_A} \in \mathcal{S}_A$ are pure and distinguishable, and $\varphi^B_1, \dots, \varphi^B_{c_B} \in \mathcal{S}_B$ are pure and distinguishable, too. This and equation (16) imply

$$\mu_{AB} = \mu_A \otimes \mu_B = \frac{1}{c_A c_B}\sum_{a=1}^{c_A}\sum_{b=1}^{c_B} \varphi^A_a \otimes \varphi^B_b \ . \tag{19}$$

All states $\varphi^A_a \otimes \varphi^B_b \in \mathcal{S}_{AB}$ are distinguishable with the tensor-product measurement, therefore

$$c_{AB} \geq c_A c_B \ . \tag{20}$$

Let $(\Omega_1, \dots, \Omega_{c_{AB}})$ be a complete measurement on AB which distinguishes the states $\psi_1, \dots, \psi_{c_{AB}} \in \mathcal{S}_{AB}$; that is $\Omega_k(\psi_{k'}) = \delta_{k,k'}$. According to Lemma 4 these states can be chosen to be pure. Since $\sum_{k=1}^{c_{AB}} \Omega_k(\mu_{AB}) = 1$, there is at least one value of k, denoted k_0, such that

$$\Omega_{k_0}(\mu_{AB}) \leq 1/c_{AB} \ . \tag{21}$$

The product of pure states $\varphi^A_1 \otimes \varphi^B_1$ is pure [3], hence Requirement 4 tells that there is a reversible transformation $G \in \mathcal{G}_{AB}$ such that $G(\psi_{k_0}) = \varphi^A_1 \otimes \varphi^B_1$. The measurement $(\Omega_1 \circ G^{-1}, \dots, \Omega_{c_{AB}} \circ G^{-1})$ distinguishes the states $G(\psi_1), \dots, G(\psi_{c_{AB}})$. Inequality (21), the invariance of μ_{AB}, expansion (19), the positivity of probabilities, and $(\Omega_{k_0} \circ G^{-1})(\varphi^A_1 \otimes \varphi^B_1) = 1$, imply

$$\begin{aligned} \frac{1}{c_{AB}} \ &\geq \ (\Omega_{k_0} \circ G^{-1})(\mu_{AB}) \\ &= \ \frac{1}{c_A c_B}\sum_{a,b} (\Omega_{k_0} \circ G^{-1})(\varphi^A_a \otimes \varphi^B_b) \geq \frac{1}{c_A c_B} \ . \end{aligned}$$

This and (20) imply (18). □

It is shown in [4] that the two multiplicativity formulas (8) and (18) imply the existence of a positive integer r such that: for any c the state space $\mathcal{S}_c$ has dimension

$$d_c = c^r - 1 \ . \tag{22}$$

The integer r is a constant of the theory, with values $r = 1$ for CPT and $r = 2$ for QT.

D. Recovering classical probability theory

Let us consider all theories with $d_2 = 1$. In this case, equation (22) becomes $d_c = c - 1$. In [4], it is shown that the only GPT with this relation between capacity and dimension is CPT, as described in Subsection II F. We reproduce the proof for completeness.

Theorem 4. *The only GPT with $d_2 = 1$ satisfying Requirements 1–5 is classical probability theory.*

Proof. Let $\mathcal{S}_c$ be a state space and $(\Omega_1, \ldots, \Omega_c)$ a complete measurement which distinguishes the states $\psi_1, \ldots, \psi_c \in \mathcal{S}_c$. The vectors $\psi_1, \ldots, \psi_c \in \mathbb{R}^c$ are linearly independent; otherwise $\psi_a = \sum_{b\neq a} t_b\, \psi_b$ and $1 = \Omega_a(\psi_a) = \sum_{b\neq a} t_b\, \Omega_a(\psi_b) = 0$ gives a contradiction. Therefore, any state $\psi \in \mathcal{S}_c \subseteq \mathbb{R}^c$ can be written in this basis $\psi = \sum_a q_a\, \psi_a$ where $q_a = \Omega_a(\psi)$ turns out to be the probability of outcome a. The numbers $(q_1, \ldots, q_c)$ constitute a probability distribution, hence, there is a one-to-one correspondence between states in $\mathcal{S}_c$ and c-outcome probability distributions. This kind of set is called a d_c-simplex. A similar argument shows that the effects $\Omega_1, \ldots, \Omega_c$ are linearly independent. Hence, any effect Ω on $\mathcal{S}_c$ can be written as $\Omega = \sum_a h_a \Omega_a$, and the constraint $0 \leq \Omega(\psi_a) \leq 1$ implies $0 \leq h_a \leq 1$. In other words, every measurement on $\mathcal{S}_c$ is generated by the complete one.

Every reversible transformation on $\mathcal{S}_c$ is a symmetry of the d_c-simplex, that is, a permutation of pure states. Due to Requirement 4, there is a reversible transformation on the bit $\mathcal{S}_2$ which exchanges the two pure states. Using Requirement 3 inductively: if there is a transformation on $\mathcal{S}_{c-1}$ which exchanges two pure states and leaves the rest invariant, this transposition can be implemented on $\mathcal{S}_c$, also leaving all other pure states invariant. Therefore, all transpositions can be implemented in $\mathcal{S}_c$, and those generate the full group of permutations. □

E. Reversible transformations for the generalized bit

In the rest of the paper, only theories with $d_2 > 1$ are considered. Theorem 2 shows that $\hat{\mathcal{S}}_2$ is a d_2-dimensional unit ball. Equation (22) for $c = 2$ implies that d_2 is odd. The pure states in $\hat{\mathcal{S}}_2$ are the unit vectors in $\mathbb{R}^{d_2}$. A reversible transformation $\hat{G} \in \hat{\mathcal{G}}_2$ maps pure states onto pure states, hence it preserves the norm and has to be an orthogonal matrix $\hat{G}^{\mathrm{T}} = \hat{G}^{-1}$. Therefore $\hat{\mathcal{G}}_2$ is a subgroup of the orthogonal group $\mathrm{O}(d_2)$.

Requirement 4 imposes that for any pair of unit vectors $\hat{\varphi}, \hat{\varphi}'$ there is $\hat{G} \in \hat{\mathcal{G}}_2$ such that $\hat{G}(\hat{\varphi}) = \hat{\varphi}'$. In other words, $\hat{\mathcal{G}}_2$ is transitive on the sphere [25, 26]. According to Lemma 6, if $\hat{\mathcal{G}}_2$ is transitive on the sphere, then the largest connected subgroup $\hat{\mathcal{C}}_2 \subseteq \hat{\mathcal{G}}_2$ is also transitive on the sphere. The matrix group $\hat{\mathcal{C}}_2$ is compact and connected, hence a Lie group (Theorem 7.31 in [21]). The classification of all connected compact Lie groups that are transitive on the sphere is done in [25, 26]. For odd d_2, the only possibility is $\hat{\mathcal{C}}_2 = \mathrm{SO}(d_2)$, except for $d_2 = 7$ where there are additional possibilities: $\hat{\mathcal{C}}_2 = M\mathbb{G}_2M^{\mathrm{T}} \subset \mathrm{SO}(7)$ for any $M \in \mathrm{O}(7)$, where $\mathbb{G}_2$ is the fundamental representation of the smallest exceptional Lie group [27]. For even d_2, there are much more possibilities [25, 26], but equation (22) implies that d_2 must be odd.

The stabilizer of the vector $\hat{\nu}_1$ defined in (14) is the subgroup $\hat{\mathcal{H}}_2 = \{\hat{G} \in \hat{\mathcal{G}}_2 : \hat{G}(\hat{\nu}_1) = \hat{\nu}_1\}$. Each transformation $\hat{H} \in \hat{\mathcal{H}}_2$ has the form

$$\hat{H} = \begin{pmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \bar{H} \end{pmatrix} ,$$

where $\bar{H} \in \bar{\mathcal{H}}_2$ is the nontrivial part. If $\hat{\mathcal{C}}_2 = \mathrm{SO}(d_2)$ then $\mathrm{SO}(d_2 - 1) \subseteq \bar{\mathcal{H}}_2$ in the fundamental representation. In the case $d_2 = 7$, if $\hat{\mathcal{C}}_2 = M\mathbb{G}_2M^{\mathrm{T}}$ then $\bar{\mathcal{H}}_2$ contains (up to the similarity M) the real 6-dimensional representation of SU(3) given by

$$\bar{H} = \begin{pmatrix} \mathrm{re}\, U & \mathrm{im}\, U \\ -\mathrm{im}\, U & \mathrm{re}\, U \end{pmatrix} , \tag{23}$$

where $\mathrm{re}\, U$ and $\mathrm{im}\, U$ are the real and imaginary parts of $U \in \mathrm{SU}(3)$ (see exercise 22.27 in [27]).

The form of $\mathcal{G}_2$ in the standard representation is not immediate, however, in the rest of the paper only the Bloch representation $\hat{\mathcal{G}}_2$ is used.

F. Two generalized bits

The joint state space of two $\mathcal{S}_2$ systems is denoted by $\mathcal{S}_{2,2}$. The multiplicativity of the capacity (18) implies that $\mathcal{S}_{2,2}$ is equivalent to $\mathcal{S}_4$. However, we write $\mathcal{S}_{2,2}$ to emphasize the bipartite structure.

In what follows, instead of using the standard representation for bipartite systems (3) we generalize the Bloch representation to two generalized bits. A state $\psi_{AB} \in \mathcal{S}_{2,2}$ has Bloch representation $\hat{\psi}_{AB} = [\alpha, \beta, C]$ with

$$\begin{aligned} \alpha^i &= 2p(x_i) - 1 \\ \beta^j &= 2p(y_j) - 1 \\ C^{ij} &= 4p(x_i, y_j) - 2p(x_i) - 2p(y_j) + 1 \end{aligned} \tag{24}$$

for $i, j = 1, \ldots, d_2$. Note that $\alpha = \hat{\psi}_A$ and $\beta = \hat{\psi}_B$ are the reduced states in the Bloch representation (13). The correlation matrix can also be written as $C^{ij} = p(x_i, y_j) - p(x_i, \bar{y}_j) - p(\bar{x}_i, y_j) + p(\bar{x}_i, \bar{y}_j)$, and characterizes the correlations between subsystems. Product states have Bloch representation

$$(\varphi_A \otimes \varphi_B)^{\wedge} = \left[\hat{\varphi}_A, \hat{\varphi}_B, \hat{\varphi}_A\hat{\varphi}_B^{\mathrm{T}}\right] , \tag{25}$$

with rank-one correlation matrix. In QT, where $d_2 = 3$, two-qubit density matrices are often represented by $[\alpha, \beta, C]$ through formula (41). Definition (24) implies

$$-1 \leq \alpha^i, \beta^j, C^{ij} \leq 1 . \tag{26}$$

The invertible map $\mathcal{L}[\psi_{AB}] = \hat{\psi}_{AB}$ defined by (24) also determines the Bloch representation of effects $\hat{\Omega} = \Omega \circ \mathcal{L}^{-1}$. In particular, the tensor-product of two effects of the form (15) is

$$(\Omega_{\varphi_A} \otimes \Omega_{\varphi_B})^\wedge[\alpha, \beta, C] = \big(1 + \hat{\varphi}_A^{\mathrm{T}}\alpha + \hat{\varphi}_B^{\mathrm{T}}\beta + \hat{\varphi}_A^{\mathrm{T}} C \hat{\varphi}_B\big)/4 \ . \tag{27}$$

The map $\mathcal{L}$ also determines the action of reversible transformation in the Bloch representation. Since $\mathcal{L}$ is affine but not linear, the action $\hat{G} = \mathcal{L} \circ G \circ \mathcal{L}^{-1}$ need not be linear. Identities (16, 25) and $\hat{\mu}_2 = \mathbf{0}$ imply that the maximally-mixed state in $\hat{\mathcal{S}}_{2,2}$ is $\hat{\mu}_{2,2} = (\hat{\mu}_2, \hat{\mu}_2, \hat{\mu}_2\hat{\mu}_2^{\mathrm{T}}) = \mathbf{0}$. This and (12) imply that transformations $\hat{\mathcal{G}}_{2,2}$ act on the generic vector $[\alpha, \beta, C]$ as matrices. In particular, local transformations $G_A, G_B \in \mathcal{G}_2$ act as

$$(G_A \otimes G_B)^\wedge[\alpha, \beta, C] = [\hat{G}_A\alpha, \hat{G}_B\beta, \hat{G}_A C \hat{G}_B^{\mathrm{T}}] \ . \tag{28}$$

Subsection IV E concludes that $\hat{\mathcal{G}}_2$ consists of orthogonal matrices, and Lemma 8 shows that all transformations in $\hat{\mathcal{G}}_{2,2}$ are orthogonal, too. Orthogonal matrices preserve the norm of vectors, therefore all pure states $\psi \in \mathcal{S}_{2,2}$ satisfy

$$|\hat{\psi}|^2 = |\alpha|^2 + |\beta|^2 + \mathrm{tr}(C^{\mathrm{T}}C) = 3 \ . \tag{29}$$

The constant in the right-hand side can be obtained by letting $\hat{\psi} = [\alpha, \alpha, \alpha\alpha^{\mathrm{T}}]$ with $|\alpha| = 1$.

G. Consistency in the subspaces of two generalized bits

In this subsection we use a trick introduced in [16]: to impose the equivalence between a particular subspace of $\mathcal{S}_{2,2}$ and $\mathcal{S}_2$ (Requirement 3).

Consider the unit vector $\hat{\nu}_1$ from (14) and the two distinguishable pure states $\hat{\varphi}_0 = \hat{\nu}_1$ and $\hat{\varphi}_1 = -\hat{\nu}_1$ from $\hat{\mathcal{S}}_2$. The four pure states $\varphi_{a,b} = \varphi_a \otimes \varphi_b \in \mathcal{S}_{2,2}$ can be distinguished with the complete measurement $\Omega_{a,b} = \Omega_{\varphi_a} \otimes \Omega_{\varphi_b}$ where $a, b \in \{0, 1\}$. Formula (27) implies

$$\hat{\Omega}_{0,0}[\alpha, \beta, C] \ = \ (1 + \alpha^1 + \beta^1 + C^{1,1})/4 \ , \tag{30}$$

$$\hat{\Omega}_{1,1}[\alpha, \beta, C] \ = \ (1 - \alpha^1 - \beta^1 + C^{1,1})/4 \ . \tag{31}$$

Requirement 3 implies that the subspace

$$\mathcal{S}_2' = \{\psi \in \mathcal{S}_{2,2} : (\Omega_{0,0} + \Omega_{1,1})(\psi) = 1\} \ ,$$

is equivalent to $\mathcal{S}_2$. By adding (30) plus (31), it becomes clear that a state $\hat{\psi} = [\alpha, \beta, C]$ belongs to $\hat{\mathcal{S}}_2'$ if and only if $C^{1,1} = 1$. Moreover, if $\psi \in \mathcal{S}_2'$, then it follows from $\hat{\Omega}_{0,1}(\hat{\psi}) \geq 0$ and $\hat{\Omega}_{1,0}(\hat{\psi}) \geq 0$ that $\alpha^1 = \beta^1$.

Theorem 5. *The state space of a generalized bit has dimension three ($d_2 = 3$).*

Proof. Recall that the case under consideration is odd d_2 larger than one. The space $\mathcal{S}_2' \subset \mathcal{S}_{2,2}$ is equivalent to $\mathcal{S}_2$, which is a d_2-dimensional unit ball. If $\varphi_{0,0}$ and $\varphi_{1,1}$ are considered the poles of this ball, then the equator is the set of states ψ_{eq} such that $\Omega_{0,0}(\psi_{\mathrm{eq}}) = \Omega_{1,1}(\psi_{\mathrm{eq}}) = 1/2$. Equations (30, 31) tell that equator states have $\alpha^1 = \beta^1 = 0$, and then

$$\hat{\psi}_{\mathrm{eq}} = [\begin{pmatrix} 0 \\ \bar{\alpha} \end{pmatrix}, \begin{pmatrix} 0 \\ \bar{\beta} \end{pmatrix}, \begin{pmatrix} 1 & \bar{\tau}^{\mathrm{T}} \\ \bar{\gamma} & \bar{C} \end{pmatrix}] \ , \tag{32}$$

where $\bar{\alpha}, \bar{\beta}, \bar{\gamma}, \bar{\tau} \in \mathbb{R}^{d_2-1}$ and $\bar{C} \in \mathbb{R}^{(d_2-1)\times(d_2-1)}$. Consider the action of $G_A \otimes \mathbb{I}$ for $G_A \in \mathcal{G}_2$ on an equator state ψ_{eq}. Since $\hat{\mathcal{G}}_2$ is transitive on the unit sphere, if $\bar{\gamma} \neq \mathbf{0}$ then there is some $\hat{G}_A \in \hat{\mathcal{G}}_2$ such that the correlation matrix transforms into

$$\hat{G}_A \begin{pmatrix} 1 & \bar{\tau}^{\mathrm{T}} \\ \bar{\gamma} & \bar{C} \end{pmatrix} = \begin{pmatrix} \sqrt{1 + |\bar{\gamma}|^2} & ? \\ \mathbf{0} & ? \end{pmatrix} \ ,$$

which is in contradiction with (26). Therefore $\bar{\gamma} = \mathbf{0}$, and by a similar argument $\bar{\tau} = \mathbf{0}$.

The stabilizer of $\hat{\nu}_1$ is the largest subgroup $\hat{\mathcal{H}}_2 \subset \hat{\mathcal{G}}_2$ which leaves $\hat{\nu}_1$ invariant (Subsection IV E). For any pair $H_A, H_B \in \mathcal{H}_2$ the identity $\Omega_{a,b} \circ (H_A \otimes H_B) = \Omega_{a,b}$ holds, which implies that if ψ_{eq} belongs to the equator (32) then

$$(H_A \otimes H_B)(\psi_{\mathrm{eq}})^\wedge = [\begin{pmatrix} 0 \\ \bar{H}_A\bar{\alpha} \end{pmatrix}, \begin{pmatrix} 0 \\ \bar{H}_B\bar{\beta} \end{pmatrix}, \begin{pmatrix} 1 & \mathbf{0}^{\mathrm{T}} \\ \mathbf{0} & \bar{H}_A\bar{C}\bar{H}_B^{\mathrm{T}} \end{pmatrix}]$$

also belongs to the equator. The equator is a unit ball of dimension $d_2 - 1$. Since the set

$$\{(H_A \otimes H_B)(\psi_{\mathrm{eq}}) : H_A, H_B \in \mathcal{H}_2\} \tag{33}$$

is a subset of the equator, the dimension of its affine span is at most $d_2 - 1$.

Consider the case $\bar{C} = \mathbf{0}$. The normalization condition (29) implies $|\bar{\alpha}| = |\bar{\beta}| = 1$. The set $\{\bar{H}_A\bar{\alpha} : \bar{H}_A \in \bar{\mathcal{H}}_2\}$ has dimension $d_2 - 1$, and the same for $\{\bar{H}_B\bar{\beta} : \bar{H}_B \in \bar{\mathcal{H}}_2\}$. Therefore the set (33) has dimension at least $2(d_2 - 1)$ generating a contradiction.

Consider the case $\bar{C} \neq \mathbf{0}$. The group action on $\bar{C}$ corresponds to the exterior tensor product $\bar{\mathcal{H}}_2 \boxtimes \bar{\mathcal{H}}_2 = \{\bar{H}_A \otimes \bar{H}_B : \bar{H}_A, \bar{H}_B \in \bar{\mathcal{H}}_2\}$. If $d_2 > 3$ and $\mathrm{SO}(d_2 - 1) \subseteq \bar{\mathcal{H}}_2$ then $\bar{\mathcal{H}}_2$ is irreducible in $\mathbb{C}^{d_2-1}$, and a simple character-based argument shows that $\bar{\mathcal{H}}_2 \boxtimes \bar{\mathcal{H}}_2$ is irreducible in $(\mathbb{C}^{d_2-1})^{\otimes 2}$ (see page 427 in [27]). Hence the set

$$\{\bar{H}_A\bar{C}\bar{H}_B^{\mathrm{T}} : \bar{H}_A, \bar{H}_B \in \bar{\mathcal{H}}_2\} \tag{34}$$

has dimension $(d_2 - 1)^2$, which conflicts with the dimensionality requirements of (33). If $d_2 = 7$ and $\bar{\mathcal{H}}_2$ contains the representation of SU(3) given in (23), then the subgroup

$$\bar{H} = \begin{pmatrix} U & \mathbf{0} \\ \mathbf{0} & U \end{pmatrix}$$

with $U \in \mathrm{SO}(3) \subset \mathrm{SU}(3)$ has two invariant $\mathbb{C}^3$ subspaces. Therefore the invariant subspaces of $\bar{\mathcal{H}}_2 \boxtimes \bar{\mathcal{H}}_2$ have at least dimension 9, and independently of $\bar{C}$, the set (34) has at least dimension 9, which conflicts with the dimensionality requirements of (33). So the only possibility is $d_2 = 3$. □

From now on, only the case $d_2 = 3$ is considered. Subsection IV E tells that $\mathrm{SO}(3) \subseteq \hat{\mathcal{G}}_2 \subseteq \mathrm{O}(3)$, which implies that either $\hat{\mathcal{G}}_2 = \mathrm{O}(3)$ and $\bar{\mathcal{H}}_2 = \mathrm{O}(2)$, or $\hat{\mathcal{G}}_2 = \mathrm{SO}(3)$ and $\bar{\mathcal{H}}_2 = \mathrm{SO}(2)$.

Let us see that the first case is impossible. The group $\bar{\mathcal{H}}_2 = \mathrm{O}(2)$ is irreducible in $\mathbb{C}^2$, therefore $\bar{\mathcal{H}}_2 \boxtimes \bar{\mathcal{H}}_2$ is irreducible in $(\mathbb{C}^2)^{\otimes 2}$. Three paragraphs above it is shown that $\bar{C} \neq \mathbf{0}$, hence the set (34) has dimension $(d_2 - 1)^2$, which is a lower bound for the one of (33), which is larger than the allowed one ($d_2 - 1 = 2$).

Let us address the second case. The group $\bar{\mathcal{H}}_2 = \mathrm{SO}(2)$ is irreducible in $\mathbb{R}^2$ but reducible in $\mathbb{C}^2$; so the previous argument does not hold. The vector space of 2×2 real matrices decomposes into the subspace generated by rotations

$$R_+ = \begin{pmatrix} \cos v & \sin v \\ -\sin v & \cos v \end{pmatrix} , \tag{35}$$

and the one generated by reflections

$$R_- = \begin{pmatrix} \cos v & \sin v \\ \sin v & -\cos v \end{pmatrix} , \tag{36}$$

where $\det R_\pm = \pm 1$. For any pair $\bar{H}_A, \bar{H}_B \in \bar{\mathcal{H}}_2$ the matrix $\bar{H}_A R_+ \bar{H}_B^{\mathrm{T}}$ is a rotation and the matrix $\bar{H}_A R_- \bar{H}_B^{\mathrm{T}}$ is a reflection; therefore the 2-dimensional subspaces (35) and (36) are invariant under $\bar{\mathcal{H}}_2 \boxtimes \bar{\mathcal{H}}_2$. Since the equator has dimension $d_2 - 1 = 2$, all matrices $\bar{C} \neq \mathbf{0}$ must be fully contained in one of the two subspaces spanned by (35) or (36), otherwise the dimension of the set (33) would be too large again. For the same reason $\bar{\alpha} = \bar{\beta} = \mathbf{0}$.

Depending on whether $\bar{C}$ is in the subspace generated by R_+ from (35) or by R_- from (36), the states in the equator of $\mathcal{S}_2'$ are either $\hat{\psi}_{\mathrm{eq}}^+$ or $\hat{\psi}_{\mathrm{eq}}^-$, where

$$\hat{\psi}_{\mathrm{eq}}^\pm = [\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos v & \sin v \\ 0 & \mp \sin v & \pm \cos v \end{pmatrix}] .$$

The proportionality constants in $\bar{C} \propto R_\pm$ are fixed by normalization (29). It turns out that both the symmetric case $\hat{\psi}_{\mathrm{eq}}^-$ and the antisymmetric case $\hat{\psi}_{\mathrm{eq}}^+$ correspond to different representations of the same physical theory —that is, the corresponding state spaces (together with measurements and transformations) are *equivalent* in the sense of Subsection II E. To see this, define the linear map $\hat{\tau} : \hat{\mathcal{S}}_2 \to \hat{\mathcal{S}}_2$ as $\hat{\tau}(\alpha_1, \alpha_2, \alpha_3)^{\mathrm{T}} := (\alpha_1, \alpha_2, -\alpha_3)^{\mathrm{T}}$; that is, a reflection in the Bloch ball. The equivalence transformation is defined as $\mathcal{L} := \tau \otimes \mathbb{I}$ (in quantum information terms, this is a "partial transposition"). This map respects the tensor product structure, leaves the set of product states invariant, and satisfies $\hat{\mathcal{L}}(\hat{\psi}_{\mathrm{eq}}^+) = \hat{\psi}_{\mathrm{eq}}^-$ [16]. In other words: we have reduced the discussion of the antisymmetric theory to that of the symmetric theory [31], which will be considered for the rest of the paper.

The orthogonality of the matrices in $\hat{\mathcal{G}}_{2,2}$ implies that $\hat{\mathcal{S}}_2'$ is a 3-dimensional ball, and not just affinely related to it. Hence all states on the surface of the ball $\hat{\mathcal{S}}_2' \subset \hat{\mathcal{S}}_{2,2}$ can be parametrized in polar coordinates $u \in [0, \pi)$ and $v \in [0, 2\pi)$ as

$$\hat{\psi}(u, v) = \tag{37}$$
$$[\begin{pmatrix} \cos u \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \cos u \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sin u \cos v & \sin u \sin v \\ 0 & \sin u \sin v & -\sin u \cos v \end{pmatrix}] .$$

These states cannot be written as proper mixtures of other states from $\hat{\mathcal{S}}_2'$. It is easy to see that this implies that they are pure states in $\hat{\mathcal{S}}_{2,2}$.

H. The Hermitian representation

In this subsection, a new (more familiar) representation is introduced, where states in $\mathcal{S}_2$ are represented by 2×2 Hermitian matrices. For any state $\psi \in \mathcal{S}_2$ in the standard representation (1), define the linear map

$$\mathcal{L}[\psi] = \psi^0 \frac{\mathbb{I} - \sigma^1 - \sigma^2 - \sigma^3}{2} + \sum_{i=1}^{3} \psi^i \sigma^i . \tag{38}$$

The Pauli matrices

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \quad \sigma^2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix} , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} ,$$

together with the identity $\mathbb{I}$ constitute an orthogonal basis for the real vector space of Hermitian matrices. In terms of the Bloch representation, the map (38) has the familiar form

$$\mathcal{L}[\psi] = \frac{1}{2} \left(\mathbb{I} + \sum_{i=1}^{3} \hat{\psi}^i \sigma^i \right) .$$

All positive unit-trace 2×2 Hermitian matrices can be written in this way with $\hat{\psi}$ in the unit sphere. Since $\hat{\mathcal{S}}_2$ is a 3-dimensional unit sphere, the set $\mathcal{L}[\mathcal{S}_2]$ is the set of quantum states. The extreme points of $\mathcal{L}[\mathcal{S}_2]$ are the rank-one projectors: each pure state $\psi \in \mathcal{S}_2$ satisfies $\mathcal{L}[\psi] = |\psi\rangle\langle\psi|$, where the vector $|\psi\rangle \in \mathbb{C}^2$ is defined up to a global phase. Effect (15) associated to the pure state $\varphi \in \mathcal{S}_2$ is

$$\Omega_\varphi(\psi) = \left(\Omega_\varphi \circ \mathcal{L}^{-1}\right)(\mathcal{L}[\psi]) = \mathrm{tr}\left(|\varphi\rangle\langle\varphi|\, \mathcal{L}[\psi]\right) . \tag{39}$$

Note that the state φ and its associated effect Ω_φ are both represented by $|\varphi\rangle\langle\varphi|$. The action of a reversible

transformation $\hat{G} \in \hat{\mathcal{G}}_2 = \mathrm{SO}(3)$ in the Hermitian representation is

$$\mathcal{L}[G(\psi)] = U\mathcal{L}[\psi]U^\dagger \,,$$

where $U \in \mathrm{SU}(2)$ is related to $\hat{G}$ via

$$\sum_{j=1}^{3} \hat{G}^{ji}\sigma^j = U\sigma^i U^\dagger \,, \tag{40}$$

and $\hat{G}^{ji}$ are the matrix components (equation VII.5.12 in [23]). In summary, the generalized bit in all theories satisfying $d_2 > 1$ and the requirements, is equivalent to the qubit in QT.

I. Reconstructing quantum theory

In this subsection, the main result of this work is proved. But before, let us introduce some notation.

In QT, the state space with capacity c and the corresponding group of reversible transformations are

$$\begin{aligned} \mathcal{S}_c^{\mathrm{Q}} &= \{\rho \in \mathbb{C}^{c\times c} :\ \rho \geq 0,\ \mathrm{tr}\rho = 1\} \,, \\ \mathcal{G}_c^{\mathrm{Q}} &= \{U \otimes U^* : U \in \mathrm{SU}(c)\} \,. \end{aligned}$$

The joint state space of m generalized bits is denoted by $\mathcal{S}_{2\times m}$, and the corresponding group of reversible transformations by $\mathcal{G}_{2\times m}$. The Hermitian representation of a state $\psi \in \mathcal{S}_{2\times m}$ is defined to be $\mathcal{L}^{\otimes m}[\psi]$, where $\mathcal{L}^{\otimes m} := \mathcal{L}\otimes\cdots\otimes\mathcal{L}$, and $\mathcal{L}$ is defined in (38). The map $\mathcal{L}^{\otimes m}$ acts independently on each tensor factor, hence it translates the tensor product structure from the standard representation (4, 5, 7) to the Hermitian one. For example, if $\varphi \in \mathcal{S}_2$ is a pure state, then $\mathcal{L}^{\otimes m}[\varphi^{\otimes m}] = |\varphi\rangle\langle\varphi|^{\otimes m}$. The notation

$$\begin{aligned} \mathcal{S}_{2\times m}^{\mathrm{H}} &= \mathcal{L}^{\otimes m}[\mathcal{S}_{2\times m}] \,, \\ \mathcal{G}_{2\times m}^{\mathrm{H}} &= \mathcal{L}^{\otimes m} \circ \mathcal{G}_{2\times m} \circ (\mathcal{L}^{\otimes m})^{-1} \,, \end{aligned}$$

will be useful. The Hermitian representation of a state $\hat{\psi}_{AB} = [\alpha, \beta, C] \in \hat{\mathcal{S}}_{2,2}$ is

$$\mathcal{L}^{\otimes 2}[\psi_{AB}] = \tag{41}$$
$$\frac{1}{4}\Big[\mathbb{I}\otimes\mathbb{I} + \sum_{i=1}^{3}\alpha^i\sigma^i\otimes\mathbb{I} + \sum_{j=1}^{3}\beta^j\mathbb{I}\otimes\sigma^j + \sum_{i,j=1}^{3} C^{ij}\sigma^i\otimes\sigma^j\Big] \,.$$

The action of local transformations $G_A, G_B \in \mathcal{G}_2$ on $\psi_{AB} \in \mathcal{S}_{2,2}$ is

$$\mathcal{L}^{\otimes 2}\big[(G_A\otimes G_B)(\psi_{AB})\big] = (U_A\otimes U_B)\rho_{AB}(U_A\otimes U_B)^\dagger \tag{42}$$

where $\rho_{AB} = \mathcal{L}^{\otimes 2}[\psi_{AB}]$ and $U_A, U_B \in \mathrm{SU}(2)$ are related to G_A, G_B via (40). Now, we are ready to prove

Theorem 6. *The only GPT with $d_2 > 1$ satisfying Requirements 1–5 is quantum theory.*

Proof. We start by reproducing an argument from [16] which shows that $\mathcal{S}_4^{\mathrm{Q}} \subseteq \mathcal{S}_{2,2}^{\mathrm{H}}$. A particular family of pure states in $\mathcal{S}_{2,2}$ is $\psi(u) = \psi(u,0)$ defined in (37). The Hermitian representation of $\psi(u)$ is the projector $\mathcal{L}^{\otimes 2}[\psi(u)] = |\psi(u)\rangle\langle\psi(u)|$ onto the $\mathbb{C}^2\otimes\mathbb{C}^2$-vector

$$|\psi(u)\rangle = \cos\frac{u}{2}|+\rangle\otimes|+\rangle + \sin\frac{u}{2}|-\rangle\otimes|-\rangle \,,$$

where $|+\rangle = (1,1)^{\mathrm{T}}/\sqrt{2}$ and $|-\rangle = (-1,1)^{\mathrm{T}}/\sqrt{2}$. From the Schmidt decomposition, it follows that all rank-one projectors in $\mathbb{C}^{4\times 4}$ can be written as $(U_A\otimes U_B)|\psi(u)\rangle\langle\psi(u)|(U_A\otimes U_B)^\dagger$ for some value of u and some local unitaries $U_A, U_B \in \mathrm{SU}(2)$. Thus, all rank-one projectors are pure states in $\mathcal{S}_{2,2}^{\mathrm{H}}$. Their mixtures generate all of $\mathcal{S}_4^{\mathrm{Q}}$, therefore $\mathcal{S}_4^{\mathrm{Q}} \subseteq \mathcal{S}_{2,2}^{\mathrm{H}}$.

Direct calculation shows that

$$\mathrm{tr}\big(\mathcal{L}^{\otimes 2}[\psi]\,\mathcal{L}^{\otimes 2}[\psi']\big) = \frac{1}{4} + \frac{1}{4}\left[\alpha^{\mathrm{T}}\alpha' + \beta^{\mathrm{T}}\beta' + \mathrm{tr}(C^{\mathrm{T}}C')\right] \tag{43}$$

for any pair of states $\hat{\psi} = [\alpha, \beta, C]$ and $\hat{\psi}' = [\alpha', \beta', C']$ from $\hat{\mathcal{S}}_{2,2}$. Lemma 8 shows that all $\hat{G} \in \hat{\mathcal{G}}_{2,2}$ are orthogonal matrices. Therefore, the Euclidean inner product between states, as in the right-hand side of (43), is preserved by the action of any $\hat{G} \in \hat{\mathcal{G}}_{2,2}$. Equality (43) maps this property to the Hermitian representation: any $H \in \mathcal{G}_{2,2}^{\mathrm{H}}$ preserves the Hilbert-Schmidt inner product between states:

$$\mathrm{tr}[H(\rho)\,H(\rho')] = \mathrm{tr}(\rho\,\rho') \,, \tag{44}$$

for all $\rho, \rho' \in \mathcal{S}_{2,2}^{\mathrm{H}}$.

For any pure state $\varphi \in \mathcal{S}_2$, the rank-one projector $|\varphi\rangle\langle\varphi|\otimes|\varphi\rangle\langle\varphi| = \mathcal{L}^{\otimes 2}[\varphi\otimes\varphi]$ is a pure state in $\mathcal{S}_{2,2}^{\mathrm{H}}$, and $\mathrm{tr}(|\varphi\rangle\langle\varphi|\otimes|\varphi\rangle\langle\varphi|\,\rho) = (\Omega_\varphi\otimes\Omega_\varphi)\circ(\mathcal{L}^{\otimes 2})^{-1}(\rho)$ is a measurement on $\mathcal{S}_{2,2}^{\mathrm{H}}$. Any rank-one projector $|\psi\rangle\langle\psi| \in \mathbb{C}^{4\times 4}$ is a pure state in $\mathcal{S}_{2,2}^{\mathrm{H}}$, hence there is $H \in \mathcal{G}_{2,2}^{\mathrm{H}}$ such that $H(|\varphi\rangle\langle\varphi|\otimes|\varphi\rangle\langle\varphi|) = |\psi\rangle\langle\psi|$. Composing the transformation H with the effect $\Omega_\varphi\otimes\Omega_\varphi$ generates the effect $(\Omega_\varphi\otimes\Omega_\varphi)\circ(\mathcal{L}^{\otimes 2})^{-1}\circ H^{-1}$, which maps any $\rho \in \mathcal{S}_{2,2}^{\mathrm{H}}$ to

$$\mathrm{tr}\big[|\varphi\rangle\langle\varphi|\otimes|\varphi\rangle\langle\varphi|\,H^{-1}(\rho)\big] = \mathrm{tr}[|\psi\rangle\langle\psi|\,\rho] \,, \tag{45}$$

where (44) has been used. In summary, every rank-one projector $|\psi\rangle\langle\psi| \in \mathbb{C}^{4\times 4}$ has an associated effect (45) which is an allowed measurement on $\mathcal{S}_{2,2}^{\mathrm{H}}$, and these generate all quantum effects.

We have seen that all quantum states $\mathcal{S}_4^{\mathrm{Q}}$ are contained in $\mathcal{S}_4^{\mathrm{H}}$, but can there be other states? If so, the associated Hermitian matrices should have a negative eigenvalue (note that all states in the Hermitian representation (41) have unit trace). If ρ has a negative eigenvalue and $|\psi\rangle$ is the corresponding eigenvector, then the associated measurement outcome (45) has negative probability. Hence, we conclude that $\mathcal{S}_4^{\mathrm{Q}} = \mathcal{S}_4^{\mathrm{H}}$, and similarly for the measurements.

All reversible transformations $H \in \mathcal{G}_4^{\mathrm{H}}$ map pure states to pure states, that is, rank-one projectors to rank-one

projectors. According to Wigner's Theorem [28], every map of this kind can be written as $H(|\psi\rangle\langle\psi|) = (U|\psi\rangle)(U|\psi\rangle)^\dagger$, where U is either unitary or anti-unitary. If U is anti-unitary, it follows from Wigner's normal form [29] that there is a two-dimensional U-invariant subspace spanned by two orthonormal vectors $|\theta_0\rangle, |\theta_1\rangle \in \mathbb{C}^4$ such that $U\left(t_0|\theta_0\rangle + t_1|\theta_1\rangle\right)$ equals either $\bar{t}_0|\theta_0\rangle + \bar{t}_1|\theta_1\rangle$ or $\bar{t}_1 e^{is}|\theta_0\rangle + \bar{t}_0 e^{-is}|\theta_1\rangle$ for some $s \in \mathbb{R}$. In both cases, U acts as a reflection in the corresponding Bloch ball, which contradicts Requirement 3 because we know that $\mathcal{G}_2 = SO(3)$. Therefore $\mathcal{G}_4^{\mathrm{H}} \subseteq \mathcal{G}_4^{\mathrm{Q}}$.

We know that $\mathcal{G}_{2,2}^{\mathrm{H}}$ contains all local unitaries. Since this group is transitive on the pure states, it contains at least one unitary which maps a product state to an entangled state. It is well-known [30] that this implies that the corresponding group of unitaries constitutes a universal gate set for quantum computation; that is, it generates every unitary operation on 2 qubits. This proves that $\mathcal{G}_4^{\mathrm{H}} = \mathcal{G}_4^{\mathrm{Q}}$.

Consider m generalized bits as a composite system. From the previously discussed case of $\mathcal{S}_{2,2}$, we know that every unitary operation on every pair of generalized bits is an allowed transformation on $\mathcal{S}_{2\times m}^{\mathrm{H}}$. But two-qubit unitaries generate all unitary transformations [30], hence $\mathcal{G}_{2^m}^{\mathrm{Q}} \subseteq \mathcal{G}_{2\times m}^{\mathrm{H}}$. By applying all these unitaries to $|\varphi\rangle\langle\varphi|^{\otimes m}$, all pure quantum states are generated, hence $\mathcal{S}_{2^m}^{\mathrm{Q}} \subseteq \mathcal{S}_{2\times m}^{\mathrm{H}}$. Reasoning as in the $\mathcal{S}_{2,2}$ case, for every rank-one projector $|\psi\rangle\langle\psi|$ acting on $\mathbb{C}^{2^m}$, the associated effect which maps $\rho \in \mathcal{S}_{2\times m}^{\mathrm{H}}$ to $\mathrm{tr}(|\psi\rangle\langle\psi|\rho)$ is an allowed measurement outcome on $\mathcal{S}_{2\times m}^{\mathrm{H}}$. This implies that all matrices in $\mathcal{S}_{2\times m}^{\mathrm{H}}$ have positive eigenvalues, therefore $\mathcal{S}_{2\times m}^{\mathrm{H}} = \mathcal{S}_{2^m}^{\mathrm{Q}}$ and $\mathcal{G}_{2\times m}^{\mathrm{H}} = \mathcal{G}_{2^m}^{\mathrm{Q}}$.

The remaining cases of capacities c that are not powers of two are treated by applying Requirement 3, using that $\mathcal{S}_c \subset \mathcal{S}_{2^m}$ for large enough m. □

V. CONCLUSION

We have imposed five physical requirements on the framework of generalized probabilistic theories. These requirements are simple and have a clear physical meaning in terms of basic operational procedures. It is shown that the only theories compatible with them are CPT and QT. If Requirement 4 is strengthened by imposing the continuity of reversible transformations, then the only theory that survives is QT. Any other theory violates at least one of the requirements, hence the relaxation of each one constitutes a different way to go beyond QT.

By using the notion of "idealized theory" (in which all mathematically well-defined measurements are allowed [22]) Requirement 5 becomes unnecessary. In this case the claim of this work is reformulated as: *Any theory which has an idealized version of itself satisfying Requirements 1, 2, 3 and 4 is embedded in QT.*

We have only derived QT for finite-dimensional Hilbert spaces, but this does not seem to be a severe restriction: It seems that all infinite-dimensional models that are actually used in physics can be constructed from the finite-dimensional theory by an appropriate approximation or limit procedure.

The standard formulation of QT includes two postulates which do not follow from our requirements: (i) the update rule for the state after a measurement, and (ii) the Schrödinger equation. If desired, these can be incorporated in our derivation of QT by imposing the following two extra requirements: (i) *if a system is measured twice "in rapid succession" with the same measurement, the same outcome is obtained both times* [4], and (ii) *closed systems evolve reversibly and continuously in time.*

This derivation of QT contains two steps which deserve a special mention. First, a direct consequence of Requirement 3 is that $\mathcal{S}_2$ is fully surrounded by pure states, which together with Requirement 4 implies that $\mathcal{S}_2$ is a ball. Second, this ball has dimension three, since the only connected Lie group transitive on the boundary of a ball, with stabilizer subgroups ($H\alpha = \alpha$) satisfying $H \boxtimes H = H \oplus \cdots$, is SO(3).

At this point, some questions are unavoidable. Can any of the Requirements be weakened? The framework of generalized probabilistic theories assumes the existence of a classical level; can this be questioned?

Acknowledgments

The authors are grateful to Anne Beyreuther, Jens Eisert, Volkher Scholz, Tony Short, and Christopher Witte for discussions. Special thanks to Lucien Hardy for pointing out the multiplicity of groups that are transitive on the sphere and, correspondingly, the need to address the 7-dimensional ball as a special case. Lluís Masanes is financially supported by Caixa Manresa, and benefits from the Spanish MEC project TOQATA (FIS2008-00784) and QOIT (Consolider Ingenio 2010), EU Integrated Project SCALA and STREP project NAMEQUAM.

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] P. W. Shor, *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer*; SIAM J. Comput. **26**(5), 1484-1509 (1997), quant-ph/9508027v2.
[3] L. Hardy, *Foliable Operational Structures for General Probabilistic Theories*; arXiv:0912.4740v1.
[4] L. Hardy; *Quantum Theory From Five Reasonable Axioms*, quant-ph/0101012v4.
[5] H. Barnum, A. Wilce, *Information processing in convex operational theories*, DCM/QPL (Oxford University 2008), arXiv:0908.2352v1.

[6] J. Barrett, *Information processing in generalized probabilistic theories*, Phys. Rev. A **75**, 032304 (2007), arXiv:quant-ph/0508211v3.

[7] G. W. Mackey; *The mathematical foundations of quantum mechanics*, (W. A. Benjamin Inc, New York, 1963).

[8] G. Birkhoff, J. von Neumann, *The Logic of Quantum Mechanics*, Annals of Mathematics, **37**, 823 (1936).

[9] G. Chiribella, G. M. D'Ariano, P. Perinotti; *Probabilistic theories with purification*; Phys. Rev. A **81**, 062348 (2010), arXiv:0908.1583v5.

[10] W. van Dam, *Implausible Consequences of Superstrong Nonlocality*, arXiv:quant-ph/0501159v1.

[11] M. Pawlowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, M. Zukowski, *A new physical principle: Information Causality*, Nature **461**, 1101 (2009), arXiv:0905.2292v3.

[12] S. Popescu, D. Rohrlich, *Causality and Nonlocality as Axioms for Quantum Mechanics*, Proceedings of the Symposium on Causality and Locality in Modern Physics and Astronomy (York University, Toronto, 1997), arXiv:quant-ph/9709026v2.

[13] M. Navascues, H. Wunderlich, *A glance beyond the quantum model*, Proc. Roy. Soc. Lond. A **466**, 881-890 (2009), arXiv:0907.0372v1.

[14] D. Gross, M. Müller, R. Colbeck, O. C. O. Dahlsten, *All reversible dynamics in maximally non-local theories are trivial*, Phys. Rev. Lett. **104**, 080402 (2010), arXiv:0910.1840v2.

[15] S. Aaronson, *Is Quantum Mechanics An Island In Theoryspace?*, quant-ph/0401062v2.

[16] B. Dakić, C. Brukner, *Quantum Theory and Beyond: Is Entanglement Special?*, arXiv:0911.0695v1.

[17] C. A. Fuchs, *Quantum Mechanics as Quantum Information (and only a little more)*, Quantum Theory: Reconstruction of Foundations, A. Khrenikov (ed.), Växjo University Press (2002), arXiv:quant-ph/0205039v1.

[18] G. Brassard, *Is information the key?* Nature Physics **1**, 2 (2005).

[19] E. M. Alfsen and F. W. Shultz, *Geometry of state spaces of operator algebras*, Birkhäuser, Boston (2003).

[20] R. T. Rockafellar, *Convex Analysis*, Princeton University Press (1970).

[21] A. Baker, *Matrix Groups, An Introduction to Lie Group Theory*, Springer-Verlag London Limited (2006).

[22] A. J. Short, *private communication.*

[23] B. Simon, *Representations of Finite and Compact Groups*, Graduate Studies in Mathematics, vol. 10, American Mathematical Society (1996).

[24] S. Straszewicz, *Über exponierte Punkte abgeschlossener Punktmengen*, Fund. Math. **24**, 139-143 (1935).

[25] A. L. Onishchik and V. V. Gorbatsevich, *Lie groups and Lie algebras I*, Encyclopedia of Mathematical Sciences 20, Springer Verlag Berlin, Heidelberg (1993).

[26] A. L. Onishchik, *Transitive compact transformation groups*, Mat. Sb. (N.S.) **60**(102):4 447–485 (1963); English translation: Amer. Math. Soc. Transl. (2) **55**, 153–194 (1966).

[27] W. Fulton, J. Harris, *Representation Theory*, Graduate texts in mathematics, Springer (2004).

[28] V. Bargmann, *Note on Wigner's Theorem on Symmetry Operations*, J. Math. Phys. **5**, 862–868 (1964).

[29] E. P. Wigner, *Normal Form of Antiunitary Operators*, J. Math. Phys. **1**, 409–413 (1960).

[30] M. J. Bremner, C. M. Dawson, J. L. Dodd, A. Gilchrist, A. W. Harrow, D. Mortimer, M. A. Nielsen, T. J. Osborne, *Practical scheme for quantum computation with any two-qubit entangling gate*, Phys. Rev. Lett. **89**:247902 (2002), arXiv:quant-ph/0207072v1.

[31] As a physical interpretation of the antisymmetric case, consider two observers who have never met before, but who have independently built devices to measure spin-$\frac{1}{2}$ particles in three orthogonal directions. If they never had the chance to agree on a common "handedness" of spatial coordinate systems, and happen to have chosen two different orientations, they will measure antisymmetric correlation matrices on shared quantum states. The "three-bit nogo result" from [16] can be interpreted as follows: if there is a third observer, then it is impossible that *every* pair of parties measures antisymmetric correlation matrices.

Appendix A: Lemmas

Lemma 1. *In any state space $\mathcal{S}_c$, the only state $\psi \in \mathcal{S}_c$ which is invariant under all reversible transformations*

$$G(\psi) = \psi \quad \textit{for all} \quad G \in \mathcal{G}_c \ , \tag{A1}$$

is the maximally-mixed state μ_c, defined in (11).

Proof. Suppose $\psi \in \mathcal{S}_c$ satisfies (A1). Any state can be written as a mixture of pure states: $\psi = \sum_k q_k \, \psi_k$. Normalization $\int_{\mathcal{G}_c} dG = 1$, condition (A1), the linearity of G, the purity of all ψ_k, the definition of μ_c, and $\sum_k q_k = 1$, imply

$$\begin{aligned} \psi \ &= \ \int_{\mathcal{G}_c} \psi \, dG = \int_{\mathcal{G}_c} G(\psi) \, dG \\ &= \ \sum_k q_k \int_{\mathcal{G}_c} G(\psi_k) \, dG = \sum_k q_k \, \mu_c = \mu_c \ , \end{aligned}$$

which proves the claim. □

Lemma 2. *If $\mathcal{G}$ is a compact real matrix group, then there is a real matrix $S > 0$ such that for each $G \in \mathcal{G}$ the matrix SGS^{-1} is orthogonal.*

Proof. Since the group $\mathcal{G}$ is compact, there is an invariant Haar measure [23], which allows to define

$$P = \int_{\mathcal{G}} G^{\mathrm{T}} G \, dG \ .$$

Since each G is invertible, the matrix $G^{\mathrm{T}}G$ is strictly positive, and P too. Define $S = \sqrt{P} > 0$ where both S, S^{-1} are real and symmetric. For any $G \in \mathcal{G}$ we have $(SGS^{-1})^{\mathrm{T}}(SGS^{-1}) = \mathbb{I}$, which implies orthogonality. □

Lemma 3. *If μ_A and μ_B are the maximally-mixed states of the state spaces $\mathcal{S}_A$ and $\mathcal{S}_B$, then the maximally-mixed state of the composite system $\mathcal{S}_{AB}$ is*

$$\mu_{AB} = \mu_A \otimes \mu_B \ .$$

Proof. The pure states ψ^A in $\mathcal{S}_A$ linearly span $\mathbb{R}^{d_A+1}$, and the pure states ψ^B in $\mathcal{S}_B$ linearly span $\mathbb{R}^{d_B+1}$. Therefore, pure product states $\psi^A \otimes \psi^B$ span $\mathbb{R}^{d_A+1} \otimes \mathbb{R}^{d_B+1}$. In particular, the maximally-mixed state (11) of $\mathcal{S}_{AB}$ can be written as

$$\mu_{AB} = \sum_{a,b} t_{a,b}\, \psi_a^A \otimes \psi_b^B \ , \tag{A2}$$

where $t_{a,b} \in \mathbb{R}$ are not necessarily positive coefficients, and all ψ_a^A, ψ_b^B are pure. From definition (1), the first component of the vector equality (A2) implies $\sum_{a,b} t_{a,b} = 1$. The maximally-mixed state is invariant under all reversible transformations, in particular the local ones

$$\begin{aligned} \mu_{AB} &= \int_{\mathcal{G}_A} dG_A \int_{\mathcal{G}_B} dG_B\, (G_A \otimes G_B)(\mu_{AB}) \\ &= \sum_{a,b} t_{a,b} \left[\int_{\mathcal{G}_A} dG_A\, G_A(\psi_a^A)\right] \otimes \left[\int_{\mathcal{G}_B} dG_B\, G_B(\psi_b^B)\right] \\ &= \sum_{a,b} t_{a,b}\, \mu_A \otimes \mu_B = \mu_A \otimes \mu_B \ , \end{aligned}$$

where the same tricks from Lemma 1 have been used. □

Lemma 4. *For every tight effect Ω, there is a pure state ψ such that $\Omega(\psi) = 1$. Also, if a measurement $\Omega_1, \ldots, \Omega_n$ distinguishes n states $\psi_1, \ldots, \psi_n$, then these states can be chosen pure.*

Proof. By definition, for each tight effect Ω there is a (not necessarily pure) state ψ' such that $\Omega(\psi') = 1$. Every ψ' can be written as a mixture of pure states ψ_k, that is $\psi' = \sum_k q_k\, \psi_k$ with $q_k > 0$ and $\sum_k q_k = 1$. Effects are linear functions such that $\Omega(\psi) \leq 1$ for any state ψ. Therefore, it must happen that all pure states ψ_k in the above decomposition satisfy $\Omega(\psi_k) = 1$.

To prove the second part, let $\psi'_1, \ldots, \psi'_n$ be the states that are distinguished by the measurement, that is $\Omega_a(\psi'_b) = \delta_{a,b}$. Every ψ'_b can be written as a convex combination of pure states $\psi'_b = \sum_k q_k\, \psi_{b,k}$. But effects are linear functions such that $0 \leq \Omega(\psi) \leq 1$ for any state ψ. Hence $\Omega(\psi'_b) = 0$ is only possible if $\Omega(\psi_{b,k}) = 0$ for all k, and similarly for the case $\Omega(\psi'_b) = 1$. It follows that $\Omega_a(\psi_{b,1}) = \delta_{a,b}$. □

Lemma 5. *If $\mathcal{S}_c$ is a state space with capacity $c \geq 1$ and μ_c the corresponding maximally-mixed state, then there are c pure distinguishable states $\psi_1, \ldots, \psi_c \in \mathcal{S}_c$ such that*

$$\mu_c = \frac{1}{c} \sum_{a=1}^{c} \psi_a \ .$$

Proof. Since $\mathcal{S}_1$ contains a single state the claim is trivially true for $c = 1$. Since $\mathcal{S}_2$ is the d_2-dimensional unit ball, two antipodal points $\hat{\varphi}_1$ and $\hat{\varphi}_2 = -\hat{\varphi}_1$ are pure, distinguishable and satisfy

$$\mu_2 = \frac{1}{2}(\varphi_1 + \varphi_2) \ . \tag{A3}$$

Now, consider the joint state space of n generalized bits, denoted $\mathcal{S}_{2\times n}$. Lemma 3 and (A3) imply that the maximally-mixed state of $\mathcal{S}_{2\times n}$ is

$$\mu_{(n)} = (\mu_2)^{\otimes n} = \frac{1}{2^n} \sum_{a_i \in \{1,2\}} \varphi_{a_1} \otimes \cdots \otimes \varphi_{a_n} \ . \tag{A4}$$

The states $\varphi_{a_1} \otimes \cdots \otimes \varphi_{a_n} \in \mathcal{S}_{2\times n}$ for all $a_1, \ldots, a_n \in \{1,2\}$ are perfectly distinguishable by the corresponding product measurement, hence the capacity of $\mathcal{S}_{2\times n}$, denoted c_n, satisfies

$$c_n \geq 2^n \ . \tag{A5}$$

Let $(\Omega_1, \ldots, \Omega_{c_n})$ be a complete measurement which distinguishes the states $\psi_1, \ldots, \psi_{c_n} \in \mathcal{S}_{2\times n}$. According to Lemma 4 these states can be chosen to be pure. Since $\sum_{k=1}^{c_n} \Omega_k(\mu_{(n)}) = 1$, there is at least one value of k, denoted k_0, such that

$$\Omega_{k_0}(\mu_{(n)}) \leq 1/c_n \ . \tag{A6}$$

The state $\varphi_1 \in \mathcal{S}_2$ from (A3) is pure, hence $\varphi_1^{\otimes n} \in \mathcal{S}_{2\times n}$ is pure too. Requirement 4 tells that there is a reversible transformation G acting on $\mathcal{S}_{2\times n}$ such that $G(\psi_{k_0}) = \varphi_1^{\otimes n}$. The measurement $(\Omega_1 \circ G^{-1}, \ldots, \Omega_{c_n} \circ G^{-1})$ distinguishes the states $G(\psi_1), \ldots, G(\psi_{c_n})$. Inequality (A6), the invariance of $\mu_{(n)}$ under G, expansion (A4), the positivity of probabilities, and $(\Omega_{k_0} \circ G^{-1})(\varphi_1^{\otimes n}) = 1$, imply

$$\begin{aligned} \frac{1}{c_n} &\geq (\Omega_{k_0} \circ G^{-1})(\mu_{(n)}) \\ &= \frac{1}{2^n} \sum_{a_i \in \{1,2\}} (\Omega_{k_0} \circ G^{-1})(\varphi_{a_1} \otimes \cdots \otimes \varphi_{a_n}) \geq \frac{1}{2^n} \ . \end{aligned}$$

This and (A5) imply $c_n = 2^n$. This together with (A4) shows the assertion of the lemma for state spaces whose capacity is a power of two. The rest of cases are shown by induction.

Let us prove that if the claim of the lemma holds for a state space with capacity c, with $c > 1$, then it holds for a state space with capacity $c - 1$ too. The induction hypothesis tells that there is a complete measurement $(\Omega_1, \ldots, \Omega_c)$ which distinguishes the pure states $\psi_1, \ldots, \psi_c \in \mathcal{S}_c$, and $\mu_c = \frac{1}{c}\sum_{k=1}^{c} \psi_k \in \mathcal{S}_c$ is the corresponding maximally-mixed state. Requirement 3 tells that the state space $\mathcal{S}_{c-1}$ is equivalent to

$$\mathcal{S}'_{c-1} = \{\psi \in \mathcal{S}_c \,|\, \Omega_1(\psi) + \cdots + \Omega_{c-1}(\psi) = 1\} \ .$$

According to Requirement 3, for each $G \in \mathcal{G}_{c-1}$ there is $G' \in \mathcal{G}_c$ which implements G on $\mathcal{S}'_{c-1}$. Hence $G'(\psi_k) \in \mathcal{S}'_{c-1}$ for $k = 1, \ldots, c-1$, which implies $(\Omega_c \circ G')(\psi_k) = 0$ for those k, and

$$\begin{aligned} \frac{1}{c} &= \Omega_c(\mu_c) = (\Omega_c \circ G')(\mu_c) \\ &= \frac{1}{c} \sum_{k=1}^{c} (\Omega_c \circ G')(\psi_k) = \frac{1}{c}(\Omega_c \circ G')(\psi_c) \ , \end{aligned}$$

and $(\Omega_c \circ G')(\psi_c) = 1$. Requirement 3 tells that the set

$$\mathcal{S}'_1 = \{\psi \in \mathcal{S}_c \,|\, \Omega_c(\psi) = 1\}$$

is equivalent to $\mathcal{S}_1$, which contains a single state. This and $\Omega_c(\psi_c) = 1$ imply that $G'(\psi_c) = \psi_c$ and then $G'(\mu'_{c-1}) = \mu'_{c-1}$, where we define

$$\mu'_{c-1} = \frac{1}{c-1}\sum_{k=1}^{c-1}\psi_k = \frac{c}{c-1}\mu_c - \frac{1}{c-1}\psi_c \ \in \mathcal{S}'_{c-1} \ .$$

For any $G \in \mathcal{G}_{c-1}$ the corresponding G' satisfies $G'(\mu'_{c-1}) = \mu'_{c-1}$. Due to Lemma 1 the invariant state μ'_{c-1} must be the maximally-mixed state in $\mathcal{S}'_{c-1}$, which has the claimed form, and Requirement 3 extends this to $\mathcal{S}_{c-1}$. □

Lemma 6. *Let $\mathcal{S}$ be a state space such that the subset of pure states $\mathcal{P}$ is a connected topological manifold. If the corresponding group of transformations $\mathcal{G}$ is transitive on $\mathcal{P}$, then the largest connected subgroup $\mathcal{C} \subseteq \mathcal{G}$ is transitive on $\mathcal{P}$, too.*

Proof. This proof involves basic notions of point set topology.

Since $\mathcal{G}$ is compact, it is the union of a finite number of (disjoint) connected components $\mathcal{G} = \mathcal{C}_1 \cup \cdots \cup \mathcal{C}_n$. If $n = 1$ the lemma is trivial. Let $\mathcal{C}$ be the connected component $\mathcal{C}_i$ containing the identity matrix $\mathbb{I}$, which is the largest connected subgroup of $\mathcal{G}$. Each connected component $\mathcal{C}_i$ is clopen (open and closed), compact and a coset of the group: $\mathcal{C}_i = G_i \circ \mathcal{C}$ for some $G_i \in \mathcal{G}$ [27].

Pick $\psi \in \mathcal{P}$, and consider the continuous surjective map $f : \mathcal{G} \to \mathcal{P}$, defined by $f(G) = G(\psi) \in \mathcal{P}$. Since $\mathcal{C}$ is compact $f(\mathcal{C}) \subseteq \mathcal{P}$ is compact too. Since the manifold $\mathcal{P}$ is in particular a Hausdorff space, $f(\mathcal{C})$ is closed. Consider the set $\mathcal{D} = f^{-1}(f(\mathcal{C}))$. If two group elements $G, H \in \mathcal{G}$ are in the same component, that is $G^{-1}H \in \mathcal{C}$, then $G \in \mathcal{D}$ implies $H \in \mathcal{D}$, using that $\mathcal{C}$ is a normal subgroup of $\mathcal{G}$. This implies that $\mathcal{D}$ is the union of some connected components $\mathcal{C}_i$, and so is $\mathcal{G}\backslash\mathcal{D}$. In particular, $\mathcal{G}\backslash\mathcal{D}$ is compact, thus $f(\mathcal{G}\backslash\mathcal{D})$ is compact, hence closed. Therefore, $f(\mathcal{C}) = \mathcal{P}\backslash f(\mathcal{G}\backslash\mathcal{D})$ is open. We have thus proven that $f(\mathcal{C}) \neq \emptyset$ is clopen. Since $\mathcal{P}$ is connected, it follows that $f(\mathcal{C}) = \mathcal{P}$. □

The following lemma shows that there are transformations for two generalized bits which perform the "classical" swap and the "classical" controlled-not in a particular basis. Note that these transformations do not necessarily swap other states that are not in the given basis, as in QT. However, they implement a minimal amount of reversible computational power which exceeds, for example, that of boxworld, where no controlled-not operation is possible [14].

Lemma 7. *For each pair of distinguishable states $\varphi_0, \varphi_1 \in \mathcal{S}_2$, there are transformations $G_{\rm swap}, G_{\rm cnot} \in \mathcal{G}_{2,2}$ such that*

$$G_{\rm swap}(\varphi_a \otimes \varphi_b) \ = \ \varphi_b \otimes \varphi_a \ , \tag{A7}$$
$$G_{\rm cnot}(\varphi_a \otimes \varphi_b) \ = \ \varphi_a \otimes \varphi_{a\oplus b} \ , \tag{A8}$$

for all $a, b \in \{0, 1\}$, where $\oplus$ is addition modulo 2.

Proof. Let (Ω_0, Ω_1) be the measurement which distinguishes (φ_0, φ_1), that is $\Omega_a(\varphi_b) = \delta_{a,b}$. Define $\psi_{a,b} = \varphi_a \otimes \varphi_b$ and $\Omega_{a,b} = \Omega_a \otimes \Omega_b$ for $a, b \in \{0, 1\}$. Define

$$\mathcal{S}'_3 \ = \ \{\psi \in \mathcal{S}_{2,2} \mid (\Omega_{0,1} + \Omega_{1,0} + \Omega_{1,1})(\psi) = 1\},$$
$$\mathcal{S}'_2 \ = \ \{\psi \in \mathcal{S}_{2,2} \mid (\Omega_{0,1} + \Omega_{1,0})(\psi) = 1\} \ ,$$

and note that $\mathcal{S}'_2 \subset \mathcal{S}'_3 \subset \mathcal{S}_{2,2}$. At the end of Subsection IV B, it is shown that two distinguishable states $\varphi_0, \varphi_1 \in \mathcal{S}_2$ have Bloch representation satisfying $\hat{\varphi}_0 = -\hat{\varphi}_1$. According to Requirement 4 there is $G \in \mathcal{G}_2$ such that $G(\varphi_0) = \varphi_1$, and by linearity, $\hat{G}(\hat{\varphi}_1) = -\hat{G}(\hat{\varphi}_0) = -\hat{\varphi}_1 = \hat{\varphi}_0$. Requirement 3 implies that there is a transformation $G'_{\rm swap}$ for $\mathcal{S}'_3$ such that $G'_{\rm swap}(\psi_{0,1}) = \psi_{1,0}$ and $G'_{\rm swap}(\psi_{1,0}) = \psi_{0,1}$. According to Lemma 5, the maximally-mixed state in $\mathcal{S}'_3$ can be written as $\mu'_3 = (\psi_{0,1} + \psi_{1,0} + \psi_{1,1})/3$. Equalities $G'_{\rm swap}(\mu'_3) = \mu'_3$ and $G'_{\rm swap}(\psi_{0,1} + \psi_{1,0}) = \psi_{0,1} + \psi_{1,0}$ imply that $G'_{\rm swap}(\psi_{1,1}) = \psi_{1,1}$. Using Requirement 3 again, there is a reversible transformation $G_{\rm swap} \in \mathcal{G}_{2,2}$ which implements $G'_{\rm swap}$ in the subspace $\mathcal{S}'_3$. Repeating the argument with the maximally-mixed state (now in $\mathcal{S}_{2,2}$) we conclude that $G_{\rm swap}(\psi_{1,1}) = \psi_{1,1}$, hence $G_{\rm swap}$ satisfies (A7).

The existence of $G_{\rm cnot}$ is shown similarly, by exchanging the roles of $\psi_{0,1}$ and $\psi_{1,1}$. □

Lemma 8. *Reversible transformations for two generalized bits in the Bloch representation (24) are orthogonal:*

$$\hat{\mathcal{G}}_{2,2} \subseteq \mathrm{O}(d_4) \ .$$

Proof. In Subsection IV F the Bloch representation for two generalized bits is defined, and it is argued that reversible transformations $\hat{\mathcal{G}}_{2,2}$ act on $[\alpha, \beta, C]$ as matrices. In particular, local transformations (28) are

$$(G_A \otimes G_B)^\wedge = \begin{pmatrix} \hat{G}_A & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \hat{G}_B & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \hat{G}_A \otimes \hat{G}_B \end{pmatrix}, \tag{A9}$$

where each diagonal block acts on an entry of $[\alpha, \beta, C]$, and $\hat{G}_A, \hat{G}_B \in \hat{\mathcal{G}}_2$. In Subsection IV E it is argued that $\hat{\mathcal{G}}_2 \subseteq \mathrm{O}(d_2)$, hence local transformations (A9) are orthogonal.

Lemma 2 shows the existence of a real matrix $S > 0$ such that for any $\hat{G} \in \hat{\mathcal{G}}_{2,2}$ the matrix $S\hat{G}S^{-1}$ is orthogonal. In particular

$$\left[S \cdot (G_A \otimes G_B)^\wedge \cdot S^{-1}\right]^{\mathrm{T}} \left[S \cdot (G_A \otimes G_B)^\wedge \cdot S^{-1}\right] = \mathbb{I} \ ,$$

which implies the commutation relation

$$S \cdot (G_A \otimes G_B)^\wedge = (G_A \otimes G_B)^\wedge \cdot S \tag{A10}$$

Subsection IV E concludes that d_2 is odd, and that $\mathrm{SO}(d_2) \subseteq \hat{\mathcal{G}}_2$ except when $d_2 = 7$, where $M\mathbb{G}_2M^{-1} \subseteq \hat{\mathcal{G}}_2$ and $\mathbb{G}_2$ is the fundamental representation of the smallest exceptional Lie group [27]. For $d_2 \geq 3$ these groups act irreducibly in $\mathbb{C}^{d_2}$, hence $\hat{\mathcal{G}}_2$ acts irreducibly in $\mathbb{C}^{d_2}$ too [27]. The first two diagonal blocks in (A9) are irreducible. The exterior tensor product of two irreducible representations (in $\mathbb{C}^d$) is also an irreducible representation, hence the third diagonal block in (A9) is also irreducible. This together with (A10) implies that

$$S = \begin{pmatrix} a\mathbb{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & b\mathbb{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & s\mathbb{I} \end{pmatrix}$$

for some $a, b, s > 0$ (Schur's Lemma [27]).

According to Lemma 7, for each unit vector $\alpha \in \hat{\mathcal{S}}_2$ there is a transformation $G_{\mathrm{swap}} \in \mathcal{G}_{2,2}$ such that

$$\begin{aligned} & \hat{G}_{\mathrm{swap}}[\alpha, 0, 0] \\ = \; & \hat{G}_{\mathrm{swap}} \left([\alpha, \alpha, \alpha\alpha^{\mathrm{T}}] + [\alpha, -\alpha, -\alpha\alpha^{\mathrm{T}}] \right) /2 \\ = \; & \left([\alpha, \alpha, \alpha\alpha^{\mathrm{T}}] + [-\alpha, \alpha, -\alpha\alpha^{\mathrm{T}}] \right) /2 \\ = \; & [0, \alpha, 0] \; . \end{aligned}$$

Since $S\hat{G}_{\mathrm{swap}}S^{-1}$ is orthogonal, the vectors $[\alpha, 0, 0]$ and $[0, ba^{-1}\alpha, 0]$ have the same modulus, hence $a = b$. Also, there is a transformation $G_{\mathrm{cnot}} \in \mathcal{G}_{2,2}$ such that

$$\begin{aligned} & \hat{G}_{\mathrm{cnot}}[0, 0, \alpha\alpha^{\mathrm{T}}] \\ = \; & \hat{G}_{\mathrm{cnot}} \left([\alpha, \alpha, \alpha\alpha^{\mathrm{T}}] + [-\alpha, -\alpha, \alpha\alpha^{\mathrm{T}}] \right) /2 \\ = \; & \left([\alpha, \alpha, \alpha\alpha^{\mathrm{T}}] + [-\alpha, \alpha, -\alpha\alpha^{\mathrm{T}}] \right) /2 \\ = \; & [0, \alpha, 0] \; . \end{aligned}$$

Since $S\hat{G}_{\mathrm{cnot}}S^{-1}$ is orthogonal, the vectors $[0, 0, \alpha\alpha^{\mathrm{T}}]$ and $[0, bs^{-1}\alpha, 0]$ have the same modulus, hence $s = b$. Consequently $S = a\mathbb{I}$, and the claim follows. □