A digital approach to quantum theory

Lluís Masanes,[1] Markus P. Müller,[2] Remigiusz Augusiak,[1] and David Pérez-García[3]

[1]*ICFO-Institut de Ciències Fotòniques, Mediterranean Technology Park, 08860 Castelldefels (Barcelona), Spain*
[2]*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada*
[3]*Dpto. Analisis Matemático and IMI, Universidad Complutense de Madrid, 28040 Madrid, Spain*

(Dated: August 3, 2012)

Does information play a significant role in the foundations of physics? Information is the abstraction that allows us to refer to the states of systems when we choose to ignore the systems themselves. The viability of this can be formalized by postulating the existence of an information unit such that the state of any system can be reversibly encoded in a sufficient number of such units (bits/qubits in the classical/quantum case). This property of classical and quantum theory is not true in general, so we promote it to a postulate. We derive the full structure of quantum theory from the following operational postulates: Continuous Reversibility, Tomographic Locality and Existence of an Information Unit, which includes Information Causality. This new axiomatization provides an alternative perspective from which to look at the physical content of quantum theory, and opens the possibility of modifying and generalizing it in new ways.

I. INTRODUCTION

The search for alternative axiomatizations of quantum theory (QT) is an old topic that goes back to Birkhoff and von Neumann [1–3]. More recently, initiated by Hardy's work [4], there has been a wave of contributions taking a more operational and less mathematical approach [5–8]. Each axiomatization emphasizes different definitorial aspects of QT, providing a new perspective from which to look at the physical content of the theory, improving our understanding of it and its relations to other theories (as, for instance, gravity), and potentially, revealing new applications for quantum information processing.

In mathematics it is often convenient to have alternative axiomatizations for the same object. For example, a topological space can be axiomatized in terms of its open sets, in terms of its closure operator, and in a variety of other ways. In physics, special relativity can be stated through Einstein's principles or the Minkowski space. One could say that the Minkowski space straightforwardly specifies the mathematical structure of space-time, while Einstein's principles have a more direct physical meaning. Analogously, the standard formulation of QT—in terms of Hilbert spaces and operators acting on them—straightforwardly specifies the mathematical structure of the states, dynamics and measurements, while the axiomatization presented in this work is more of Einstein's type—it imposes some physically meaningful features that the theory must satisfy.

In this work we introduce a postulate named Existence of an Information Unit, which essentially states that there is only one type of information within the theory. Consequently, any physical process can be simulated with a suitably programmed general purpose simulator. As the input and output of the simulation need not be classical, this also generalizes the Church-Turing-Deutsch Principle (stated in [9]). An alternative way way to read this is that, at some level, the dynamics of any system is substrate-independent. Existence of an Information Unit allows us to refer to states, dynamics and measurements abstractly, without specifying the type of system they pertain to. And this is exploited by quantum information scientists, who design algorithms and protocols at an abstract level, without considering whether they will be implemented with light, atoms or any other type of physical substrate.

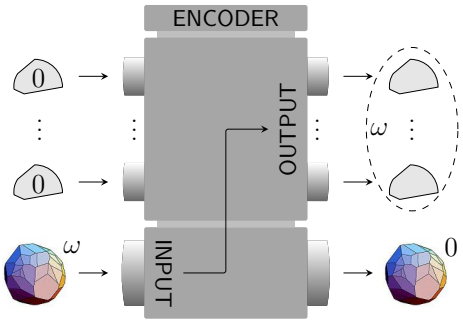

FIG. 1. *Coding* is an ideal physical transformation which maps the unknown state ω of an arbitrary system to an n-gbit state in a reversible way, and leaves the initial system in a reference state 0. Reversibility means that there is another ideal physical transformation, *decoding*, which undoes the above, bringing the arbitrary system back to its original state.

More precisely, Existence of an Information Unit states that there is a type of system, the generalized bit or gbit, such that the state of any other system can be reversibly encoded in a sufficient number of gbits (see Figure 1). In classical probability theory the gbit is the bit, and in QT it is the qubit. The reversibility of the encoding implies a correspondence between the states of any system and the states of a multi-gbit system (or an appropriate subspace). This correspondence also extends to dynamics and measurements: if our system lacks a particular dynamics then we can encode its state into a multi-gbit system, engineer the desired multi-gbit dynamics, and decode back the resulting state to our system—effectively implementing the desired dynamics in our system. We also require that gbits have some additional properties, in particular Information Causality [10].

We prove that QT is the only theory satisfying the postulates of Continuous Reversibility, Tomographic Locality (both introduced in [4] and used in many other axiomatizations of QT), and Existence of an Information Unit. As an immedi-

ate consequence, we obtain a complete characterization of the set of quantum correlations in terms of Information Causality and the above additional postulates. It is known that Information Causality on its own is not sufficient to exclude all non-quantum correlations [11], hence, our contribution also solves the problem of finding which extra assumptions suffice to completely characterize set of quantum correlations.

II. GENERAL PROBABILISTIC THEORIES

QT can be seen as a particular extension of classical probability theory, in which two random variables do not necessarily have a joint probability distribution, or in other words, cannot be jointly measured (like non-commuting observables in QT). This idea was generalized by Birkhoff and von Neumann [1], who formally defined all such extensions, nowadays referred to as general probabilistic theories (GPTs) or the convex operational framework. Recently, a lot of interest have been directed to the study of GPTs [4–8, 10, 12–23], with the double aim of constructing alternative axiomatizations of QT, and exploring what lies beyond. This, in particular, led to the discovery that many features originally thought as specific to QT (such as for instance Bell-inequality violation [22], no-cloning [14, 23], monogamy of correlations [23], Heisenberg-type uncertainty relations [17, 23], measurement-disturbance tradeoffs [14], and the possibility of secret key distribution [24, 25]), are common to most GPTs. In this light, the standard question "why does nature seem to be quantum instead of being classical?" sounds less appropriate than asking why QT instead of any other GPT. Here we answer this question by showing that any GPT different from QT violates at least one of our physically meaningful postulates. Before that let us first briefly recall the formalism of GPTs (a more detailed introduction to GPTs can be found in Appendix A).

In QT states are represented by density matrices. But, how can we represent states in theories that we do not yet know? We can follow [4]. The state of a system is represented by the probabilities of some pre-established measurement outcomes $x_1, \dots x_k$ which are called *fiducial*:

$$\omega = \begin{bmatrix} p(x_1) \\ \vdots \\ p(x_k) \end{bmatrix} \in \mathcal{S} \subset \mathbb{R}^k \,. \tag{1}$$

This list of probabilities has to be minimal but contain sufficient information to predict the probability distribution of all measurements that can be in principle performed on the system. (Note that this is always possible since the list could contain the probabilities corresponding to all measurements. In particular, the list can be infinite, that is $k = \infty$.). The number of fiducial outcomes k is equal to the dimension of $\mathcal{S}$, as otherwise one fiducial probability would be functionally related to the others, and the list not minimal. We include the possibility that the system is present with certain probability $U \in [0, 1]$ being the sum of probabilities for all the outcomes of a measurement. When the system is absent ($U = 0$) the fiducial outcomes have zero probability, hence the corresponding state (1) is the null vector $\mathbf{0} \in \mathcal{S}$. The subset of normalized states $\mathcal{N} \subset \mathcal{S}$ contains those satisfying $U(\omega) = 1$, and has dimension $k - 1$.

By the rules of probability, the set of all the allowed states $\mathcal{S}$ is convex. Indeed, by preparing the state ω_1 with probability q and ω_2 with probability $1 - q$, we effectively prepare the mixed state $q\omega_1 + (1 - q)\omega_2$. The *pure states* of $\mathcal{S}$ are the normalized states that cannot be written as mixtures. As an instance, the fiducial outcomes for a qubit can be chosen to be $\sigma_x = 1, \sigma_y = 1, \sigma_z = 1, \sigma_z = -1$, and $U(\omega) = p(\sigma_z = 1) + p(\sigma_z = -1)$. Note, however, that the set of fiducial outcomes need not be unique, nor simultaneously measurable.

In the formalism of GPTs every convex set can be seen as the state space $\mathcal{S}$ of an imaginary type of system, which, in turn, allows for construction of multipartite states spaces which violate Bell inequalities more (or less) than QT. This illustrates the degree to which this formalism generalizes classical probability theory and QT, and allows us to catch a glimpse on the multitude of alternative theories that we are considering here.

The probability of the measurement outcome x when the system is in the state ω is given by $E_x(\omega)$, where $E_x : \mathbb{R}^k \to \mathbb{R}$ is a linear function satisfying $E_x(\mathcal{S}) \subseteq [0, 1]$. To see this, suppose the system is prepared in the mixture $q\omega_1 + (1 - q)\omega_2$. Then the relative frequency of an outcome x should not depend on whether the label of the actual preparation ω_k is ignored before or after the measurement. As a result

$$E_x\big(q\omega_1 + (1 - q)\omega_2\big) = qE_x(\omega_1) + (1 - q)E_x(\omega_2) \,,$$

which together with $E_x(\mathbf{0}) = 0$ imply the linearity of E_x. In classical probability theory and QT, all such linear functions correspond to outcomes of measurements. Although this need not be the case in general, below we postulate it to hold for gbits.

Physical systems evolve with time. Often, the dynamics of a system can be controlled by adjusting its environment, allowing in this way to engineer different transformations of the system. A transformation can be represented by a map $T : \mathcal{S} \to \mathcal{S}$ which, for the same reason as outcome probabilities E, has to be linear. Sometimes there are pairs of transformations whose composition leaves the system unaffected, independently of its initial state—in this case we say that these transformations are reversible. The set of reversible transformations generated by time-continuous dynamics forms a compact connected Lie group $\mathcal{G}$. Then, the elements of the corresponding Lie algebra are the Hamiltonians of the theory (which in general have nothing to do with Hermitian matrices). Our first postulate imposes that this set of Hamiltonians is sufficiently rich.

III. THE POSTULATES

Now we are ready to present our new axiomatization of QT (see Appendix B for more details). The first postulate is motivated by the fact that most fundamental theories that we know

(classical or quantum) enjoy time-continuous reversible dynamics.

Postulate 1 (Continuous Reversibility). In any system, for every pair of pure states one can in principle engineer a time-continuous reversible dynamics which brings one state to the other.

As pointed out by Hardy [4], classical probability theory violates this postulate, since the set of reversible transformations is the group of permutations, which is discrete. This may seem contradictory with the continuity of time evolution in classical mechanics, but these are two different notions of continuity. In classical mechanics, the evolution of a pure state for an arbitrarily small time produces a new state which is perfectly distinguishable from the previous one. Hence, its dynamics is not continuous in our sense. If we relax this continuity requirement then classical probability theory also satisfies our postulates, but other theories may do as well.

Let now A and B be two systems with fiducial outcomes $x_1, \dots x_{k_A}$ and $y_1, \dots y_{k_B}$, respectively. Is there any relation between these and the fiducial outcomes of the composite system AB? The following postulate implies that the set of joint outcomes (x_i, y_j) for all i, j is a fiducial set for the composite system. As a consequence, joint probabilities (and similarly joint transformations) can be obtained through the simple tensor-product rule $p(x, y) = (E_x \otimes E_y)(\omega_{AB})$, where

$$\omega_{AB} = \begin{bmatrix} p(x_1, y_1) \\ p(x_1, y_2) \\ \vdots \\ p(x_{k_A}, y_{k_B}) \end{bmatrix} \in \mathcal{S}_{AB} \subset \mathbb{R}^{k_A} \otimes \mathbb{R}^{k_B} .$$

This also implies the multiplicity of dimensions: $k_{AB} = k_A k_B$.

Postulate 2 (Tomographic Locality). The state of a composite system is completely characterized by the correlations of measurements on the individual components.

The third postulate, introduced for the first time in this work, states the aforementioned existence of the gbit and imposes four properties that it must satisfy.

Postulate 3 (Existence of an Information Unit). There is a type of system (the gbit) such that the state of any system can be reversibly encoded in a sufficiently large number of gbits. Additionally, gbits satisfy the following:

1. State tomography is possible: $k_{\rm gbit} < \infty$.
2. All effects are observable. All linear functions $E : \mathcal{S}_{\rm gbit} \to [0, 1]$ correspond to outcomes of measurements.
3. Gbits can interact. The group of time-continuous reversible transformations for two gbits contains an element which is not product ($G_{AB} \neq G_A \otimes G_B$).
4. Weak Information Causality: when a gbit is being used to perfectly encode a classical bit, it cannot simultaneously encode any other information.

Let us explain with more detail the content of Postulate 3. First, the requirement that the state of any system can be reversibly encoded in a number of gbits means precisely that to any state space $\mathcal{S}$ (with dimension k) one can always associate a number n and a physical transformation T mapping reversibly $\mathcal{S}$ to the state space of n gbits $\mathcal{S}^n_{\rm gbit}$. More formally, there is an injective linear map $T : \mathbb{R}^k \to \mathbb{R}^{k^n_{\rm gbit}}$ satisfying the consistency constraints $T(\mathcal{S}) \subseteq \mathcal{S}^n_{\rm gbit}$ and $T^{-1}(\mathcal{S}^n_{\rm gbit} \cap {\rm Im}T) \subseteq \mathcal{S}$, where ${\rm Im}T$ is the image of T.

Second, tomography is impossible if the number of parameters defining the state is infinite. This is so because, as a consequence of weak Information Causality, gbits have two perfectly distinguishable states and no more. Note that, in QT, infinite-dimensional state tomography is possible if bounds on the energy are assumed, because infinite-dimensional systems have an infinite number of perfectly distinguishable states. But the last is not true for gbits.

Third, interaction is fundamentally necessary in order not to have an essentially trivial Universe. The requirement that any system can be reversibly encoded in gbits implies that, if gbits do not interact among them, then no other system interacts. Postulate 3.3 rules out this possibility.

Fourth, to illustrate weak Information Causality (Postulate 3.4) let us consider a communication task involving two distant parties, Alice and Bob. Suppose Alice is given two bits $a, a' \in \{0, 1\}$, and Bob is asked to guess one of them. He will base his guess on information sent to him by Alice, encoded in one gbit. Alice encodes the gbit with no knowledge of which of the two bits, a or a', Bob will try to guess. Weak Information Causality imposes that, in a coding/decoding strategy in which Bob can guess a with probability one, he knows nothing about a'. In particular, this implies that $\mathcal{S}_{\rm gbit}$ contains at most two perfectly distinguishable states. A more formal way to state Weak Information Causality is: suppose that Alice encodes a, a' in the four states $\omega_{a,a'} \in \mathcal{N}_{\rm gbit}$. If there is an effect E such that $E(\omega_{a,a'}) = \delta_{a,0}$ then any effect E' satisfies $E'(\omega_{a,0}) = E'(\omega_{a,1})$. As it is illustrated in Figure 2, this together with "all effects are observable" (cf. Postulate 3.2) imply that all states in the boundary of $\mathcal{N}_{\rm gbit}$ are pure (first arrow in Figure 3).

Having stated our three postulated, let us now show that the only theory obeying them is QT. In what follows we present an overview of the proof, while its detailed version can be found in Appendix C. Continuous Reversibility associates to any state space $\mathcal{S}$ a group of reversible transformations $\mathcal{G}$, having an invariant scalar product with respect to which all pure states of $\mathcal{S}$ have the same norm. This together with the fact that the boundary of $\mathcal{N}_{\rm gbit}$ contains only pure states imply that it is an ellipsoid (second arrow in Figure 3). By setting as the new set of fiducial outcomes the effects corresponding to the principal axes of the ellipsoid (recall that all effects are observable), $\mathcal{N}_{\rm gbit}$ becomes a Euclidean ball (third arrow in Figure 3). But what is the state space of two gbits $\mathcal{S}^2_{\rm gbit}$? According to Continuous Reversibility the set of pure states of two gbits can be written as $\{G(\omega \otimes \omega) | G \in \mathcal{G}^2_{\rm gbit}\}$, where $\mathcal{G}^2_{\rm gbit}$ is the group of reversible transformations for two gbits, and ω is a pure state of one gbit. The group $\mathcal{G}^2_{\rm gbit}$ is unknown,

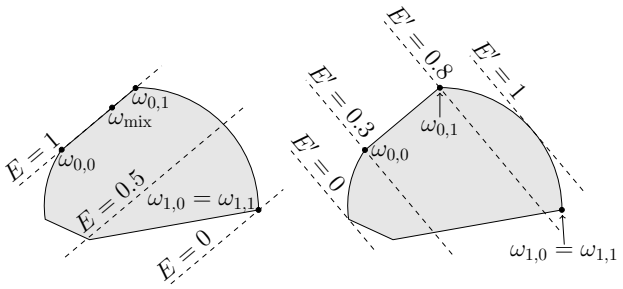

FIG. 2. This figure shows that there cannot be mixed states in the boundary of $\mathcal{N}_{\rm gbit}$. If there is one, say $\omega_{\rm mix}$, then this boundary contains a facet (left figure). Since all effects are observable, we can decode a with the effect E, which gives probability one for all states inside that facet, and probability zero for some other state(s). By encoding $(a,a')=(0,0),(0,1)$ in two different states inside that facet we can perfectly retrieve a through E, while still getting some partial information about a' with any effect E' orthogonal to the facet.

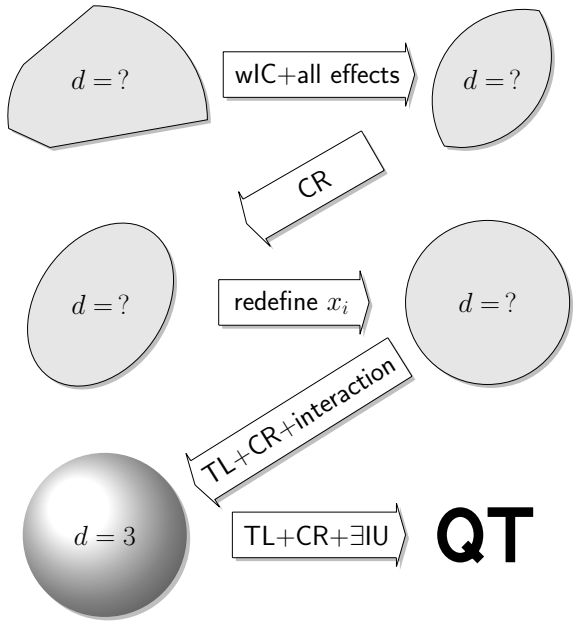

FIG. 3. This figure synthesizes the proof that the only theory satisfying the three postulates is QT. Each step (represented by an arrow) invokes part of the content of the postulates (specified inside the arrow) and reveals new information about the state space of the generalized bit. Initially (top-left) $\mathcal{N}_{\rm gbit}$ is an arbitrary convex set with arbitrary dimension $d=k_{\rm gbit}-1$, and finally (down-left) it is a 3-dimensional ball. The first arrow represents the step explained in Figure 2. The postulates of Continuous Reversibility, Tomographic Locality and Existence of an Information Unit are abbreviated by CR, TL, ∃IU.

but by consistency, it must contain all local transformations $\mathcal{G}_{\rm gbit}\otimes\mathcal{G}_{\rm gbit}$, and it must generate states with well-defined probabilities, meaning that

$$(E_x\otimes E_y)(G(\omega\otimes\omega))\in[0,1]\,, \tag{2}$$

holds for all $G\in\mathcal{G}^2_{\rm gbit}$ and any gbit effects E_x,E_y. Additionally, according to Postulate 3.4, $\mathcal{G}^2_{\rm gbit}$ must be larger than $\mathcal{G}_{\rm gbit}\otimes\mathcal{G}_{\rm gbit}$. The combination of these three requirements is very restrictive, since it implies that the Euclidean ball $\mathcal{N}_{\rm gbit}$ has dimension $k-1=3$ and $\mathcal{G}_{\rm gbit}={\rm SO}(3)$ (see Supplementary Information). This tells us that, locally, gbits are identical to qubits, but it is not clear yet whether multi-gbit state spaces $\mathcal{S}^n_{\rm gbit}$ having a non-quantum structure are consistent with our postulates. In Reference [18] all possible joint state spaces of n systems that are locally qubits are classified, and it is found that the only possibility allowing for non-product reversible transformations is multi-qubit QT. So gbits are locally and globally like qubits: $\mathcal{S}^n_{\rm gbit}$ is the set of 2^n-dimensional density matrices and $\mathcal{G}^n_{\rm gbit}$ is the adjoint representation of ${\rm SU}(2^n)$. Finally, since any state space is reversibly encodable in a multi-qubit system, the states, transformations and measurements of any system can be represented within the formalism of QT.

Conclusions.–Some attempts to go beyond QT by modifying the standard postulates [28] lead to inconsistencies [29]. On the contrary, if we stick to the framework of GPTs, no matter how we relax or modify our postulates we obtain a consistent theory. In the Supplementary Information we relax that "gbits can interact" (Postulate 3.4) and characterize the family of theories that emerge. One could also relax that "state tomography is possible" for gbits (Postulate 3.2), by considering transitive groups in the infinite-dimensional sphere. One could also relax the continuity part of Continuous Reversibility, by including the analysis of non-connected Lie groups. Hence, we believe that this axiomatization is a good starting point to modify and generalize QT, which could be a route to Quantum Gravity.

A repeated pattern in the history of science is the promotion of a scientific instrument to a model for understanding the world. For instance, there are some proposals for viewing the universe as a giant computer (classical [26] or quantum [27]). But what is the physical content of this? Can the dynamics of any system be understood as computation? After all it is computing its future state. A requisite for upgrading time-evolution to computation could be that such time-evolution is substrate-independent, so that it can be simulated in a system of information units. From this perspective, we promote the Existence of an Information Unit to be a postulate of QT, and we show that this, together with other postulates, singles out QT.

Acknowledgments.–Ll. M. acknowledges support from CatalunyaCaixa. Research at Perimeter Institute for Theoretical Physics is supported in part by the Government of Canada through NSERC and by the Province of Ontario through MRI. D. P.-G. acknowledges support from the Spanish grants I-MATH, MTM2008-01366, S2009/ESP-1594 and the European project QUEVADIS. R. A. acknowledges support from AQUTE, TOQATA, and Spanish MINCIN through the Juan de la Cierva program.

[1] G. Birkhoff and J. von Neumann, *The Logic of Quantum Mechanics*, Ann. Math. **37**, 823 (1936).

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

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

[4] L. Hardy, *Quantum theory from five reasonable axioms*, arXiv:quant-ph/0101012.

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

[6] Ll. Masanes and M. P. Müller, *A derivation of quantum theory from physical requirements*, New J. Phys. **13**, 063001 (2011); arXiv:1004.1483.

[7] L. Hardy, *Reformulating and Reconstructing Quantum Theory*, arXiv:1104.2066.

[8] G. Chiribella, G. M. D'Ariano, P. Perinotti, *Informational derivation of Quantum Theory*, Phys. Rev. A **84**, 012311 (2011); arXiv:1011.6451.

[9] D. Deutsch, *Quantum theory, the Church-Turing principle and the universal quantum computer*, Proc. R. Soc. Lond. A **400**, 97 (1985).

[10] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, *A new physical principle: Information Causality*, Nature **461**, 1101 (2009); arXiv:0905.2292v3.

[11] R. Gallego, L. E. Würflinger, A. Acín, and M. Navascués, *Quantum correlations require multipartite information principles*, Phys. Rev. Lett. **107**, 210403 (2011).

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

[13] H. Barnum, C. P. Gaebler, and A. Wilce, *Ensemble Steering, Weak Self-Duality, and the Structure of Probabilistic Theories*, arXiv:0912.5532.

[14] J. Barrett, *Information processing in generalized probabilistic theories*, arXiv:quant-ph/0508211.

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

[16] H. Barnum, J. Barrett, L. O. Clark, M. Leifer, R. Spekkens, N. Stepanik, A. Wilce, and R. Wilke, *Entropy and Information Causality in General Probabilistic Theories*, New J. Phys. **12**, 033024 (2010); arXiv:0909.5075

[17] J. Oppenheim and S. Wehner, *The uncertainty principle determines the non-locality of quantum mechanics*, Science **330**, 1072 (2010).

[18] G. de la Torre, Ll. Masanes, A. J. Short, and M. P. Müller, *Deriving quantum theory from its local structure and reversibility*, arXiv:1110.5482v1.

[19] A. J. Short and J. Barrett, *Strong nonlocality: A trade-off between states and measurements*, New J. Phys. **12**, 033034 (2010).

[20] A. J. Short and S. Wehner, *Entropy in general physical theories*, New J. Phys. **12**, 033023 (2010).

[21] C. Pfister, *One simple postulate implies that every polytopic state space is classical*, arXiv:1203.5622.

[22] S. Popescu and 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.

[23] Ll. Masanes, A. Acín, and N. Gisin, *General properties of Nonsignaling Theories*, Phys. Rev. A. **73**, 012112 (2006).

[24] J. Barrett, L. Hardy, and A. Kent, *No Signaling and Quantum Key Distribution*, Phys. Rev. Lett. **95**, 010503 (2005).

[25] Ll. Masanes, *Universally-composable privacy amplification from causality constraints*, Phys. Rev. Lett. **102**, 140501 (2009).

[26] K. Zuse, *Rechnender Raum* (Friedrich Vieweg & Sohn, Braunschweig, 1969); Tanslated as *Calculating Space*; MIT Technical Translation AZT-70-164-GEMIT, Massachusetts Institute of Technology (Project MAC), Cambridge, Mass. 02139.

[27] S. Lloyd, *Programming the Universe* (Random House, 2011).

[28] S. Weinberg, *Testing quantum mechanics*, Ann. Phys. (NY) **194**, 336 (1989).

[29] N. Gisin, *Weinberg's non-linear quantum mechanics and supraluminal communications*, Phys. Lett. A **143**, 1 (1990).

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

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

[32] L. Hardy and W. K. Wootters, *Limited Holism and Real-Vector-Space Quantum Theory*, arXiv:1005.4870.

[33] Ll. Masanes, M. P. Müller, D. Pérez-García, and R. Augusiak, *Entangling dynamics beyond quantum theory*, arXiv:1111.4060v1.

[34] B. S. Cirel'son, *Quantum generalizations of Bell's inequality*, Lett. Math. Phys. **4**, 93 (1980).

[35] J. Allcock, N. Brunner, M. Pawłowski, and V. Scarani, *Recovering part of the quantum boundary from information causality*, Phys. Rev. A **80**, 040103 (2009).

APPENDICES

Appendix A: More on general probabilistic theories

Here we recall in more detail the formalism that allows us to represent states, measurements and transformations in a theory-independent way. More complete material can be found in [7, 21].

1. States

In this formalism, the state of a system is represented by the probabilities of some pre-established measurement outcomes $x_1, \dots x_k$ which are called *fiducial*:

$$\omega = \begin{bmatrix} p(x_1) \\ \vdots \\ p(x_k) \end{bmatrix} \in \mathcal{S} \subset \mathbb{R}^k \,. \tag{A1}$$

This list of probabilities has to be minimal but contain sufficient information to predict the probability distribution of all measurements that can be in principle performed on the system. Note that this is always possible since the list could contain the probabilities corresponding to all measurements. In particular, the list can be infinite, i.e. $k = \infty$. We include the possibility that the system is present with a certain probability $u \in [0, 1]$. The value of $u = U(\omega)$ is equal to the

sum of probabilities for all the outcomes of a measurement. When the system is absent ($u = 0$) the fiducial outcomes have zero probability, hence the corresponding state (A1) is the null vector $\mathbf{0} \in \mathcal{S}$. The subset of normalized states is $\mathcal{N} = \{\omega \in \mathcal{S} | U(\omega) = 1\}$. Clearly, $\mathcal{S}$ is the convex hull of $\mathcal{N}$ and $\mathbf{0}$ [30].

By the rules of probability, the set of all the allowed states $\mathcal{S}$ is convex. Indeed, by preparing the state ω_1 with probability q and the state ω_2 with probability $1-q$, we effectively prepare the mixed state $q\omega_1 + (1-q)\omega_2$. The *pure states* of $\mathcal{S}$ are the normalized states that cannot be written as mixtures, that is, the extremal points of $\mathcal{N}$. Hence, we denote the set of pure states by $\mathrm{ext}\mathcal{N}$. The number of fiducial outcomes k is equal to the dimension of $\mathcal{S}$, as otherwise one fiducial probability would be functionally related to the others, and the list not minimal. Hence, the dimension of $\mathcal{N}$ is $k-1$. As an instance, the fiducial outcomes for a quantum two-level system (or qubit) can be chosen to be $\sigma_x = 1, \sigma_y = 1, \sigma_z = 1, \sigma_z = -1$; hence, $k = 4$ and $U(\omega) = p(\sigma_z = 1) + p(\sigma_z = -1)$. Note, however, that the set of fiducial outcomes need not be unique, nor simultaneously measurable.

By changing the set of fiducial outcomes one can transform the geometry of a state space. However, as shown in the next paragraph, all such transformations are linear and invertible. Conversely, all invertible linear transformations generate an equivalent state space, hence, state spaces are equivalence classes of convex sets under linear reversible maps. Indeed, for any invertible linear transformation $L : \mathbb{R}^k \to \mathbb{R}^k$, we can redefine the states $\omega \to L(\omega)$ and the effects $E \to E \circ L^{-1}$ such that the physics is unchanged $(E \circ L^{-1})(L(\omega)) = E(\omega)$. In a similar fashion, by redefining the transformations as $T \to L \circ T \circ L^{-1}$, the dynamical structure of the system is unchanged $(L \circ T \circ L^{-1})(L(\omega)) = L(T(\omega))$. Hence, every possible state space is an equivalence class of convex sets related by linear transformations. Note that in general, the components of the vector $L(\omega)$ are not in $[0, 1]$, so we cannot interpret them as fiducial probabilities. However, as illustrated below, sometimes it is advantageous to loose the probability interpretation of the components of $L(\omega)$ in favor of a different representation that is easier to handle.

2. Measurements

The probability of the measurement outcome x when the system is in state ω is given by $E_x(\omega)$ where $E_x : \mathbb{R}^k \to \mathbb{R}$ is a linear functional satisfying $E_x(\mathcal{S}) \subseteq [0, 1]$. To see this, suppose the system is prepared in the mixture $q\omega_1 + (1-q)\omega_2$. Then the relative frequency of an outcome x should not depend on whether the label of the actual preparation ω_k is ignored before or after the measurement. As a result

$$E_x\big(q\omega_1 + (1-q)\omega_2\big) = qE_x(\omega_1) + (1-q)E_x(\omega_2),$$

which together with $E_x(\mathbf{0}) = 0$ imply the linearity of E_x. Linear functions E satisfying $E(\mathcal{S}) \subseteq [0, 1]$ are called *effects* and can be written as a scalar product $E(\omega) = E \cdot \omega = \sum_{i=1}^{k} E^i p(x_i)$ with E being a vector from $\mathbb{R}^k$. An effect that plays a special role is $U(\omega) = \sum_{i=1}^{k} U^i p(x_i)$, which gives the probability that the system is present. In classical probability theory and QT, all effects correspond to outcomes of measurements, but this need not be the case in general (this is related to the discussion in Appendix B). Below we postulate this to hold for gbits.

An n-outcome measurement is represented by n effects $E_1, \ldots, E_n$ satisfying

$$E_1 + \cdots + E_n = U.$$

Alternatively speaking, this formula means that the outcome probabilities are normalized, implying that we only need to specify $n-1$ effects. In particular, a two-outcome measurement is represented by a single effect E, which, for a normalized state $\omega \in \mathcal{N}$, gives outcome probabilities $E(\omega)$ and $1 - E(\omega)$. We say that $\omega_1, \ldots, \omega_n \in \mathcal{S}$ are perfectly distinguishable states if there is an n-outcome measurement in $\mathcal{S}$ such that $E_i(\omega_j) = \delta_{ij}$, where δ_{ij} is the Kronecker tensor.

3. Transformations

Physical systems evolve with time. Often, the dynamics of a system can be controlled by adjusting its environment, allowing in this way to engineer different transformations of the system. A transformation is represented by a map $T : \mathbb{R}^k \to \mathbb{R}^k$ which, for the same reason as outcome probabilities E, has to be linear and satisfy the consistency constraint $T(\mathcal{S}) \subseteq \mathcal{S}$.

There can be pairs of transformations whose composition leaves the system unaffected, independently of its initial state—in this case, one of them is the inverse T^{-1} of the other, T, and we say that T is *reversible*. The set of transformations generated by time-continuous reversible dynamics forms a connected matrix group $\mathcal{G}$. From a physical point of view, it makes sense to include in $\mathcal{G}$ all transformations which can be approximated arbitrarily well by those allowed by the theory. Therefore, $\mathcal{G}$ is a compact matrix group, which according to [31], must be a Lie group. The elements of the corresponding Lie algebra are the Hamiltonians of the theory (which in general have nothing to do with Hermitian matrices). The postulate of Continuous Reversibility imposes that this set of Hamiltonians is sufficiently rich. Clearly, any reversible transformation $G \in \mathcal{G}$ must preserve the normalization of a state $U(G(\omega)) = U(\omega)$ for all $\omega \in \mathcal{S}$. This implies that $U \circ G = U$ for all $G \in \mathcal{G}$.

One can implement transformations which, in addition to a possible change of state, also transform the type of system. A transformation that takes a system from a state space $\mathcal{S}_1$ and outputs a system from a different state space $\mathcal{S}_2$, with respective dimensions k_1 and k_2, is represented by a linear map $T : \mathbb{R}^{k_1} \to \mathbb{R}^{k_2}$ satisfying the consistency constraint $T(\mathcal{S}_1) \subseteq \mathcal{S}_2$. A way to implement such a transformation is illustrated in Figure 1, where the input is an $\mathcal{S}_1$-system in an arbitrary state $\omega \in \mathcal{S}_1$ together with an $\mathcal{S}_2$-system in a fixed state 0, and the output is an $\mathcal{S}_1$-system in a fixed state 0 together with an $\mathcal{S}_2$-system in the output state $T(\omega) \in \mathcal{S}_2$.

We say that a transformation which modifies the type of system $T : \mathbb{R}^{k_1} \to \mathbb{R}^{k_2}$ is reversible if it is injective with left-inverse $T^{-1} : \mathrm{Im}T \to \mathbb{R}^{k_1}$ satisfying the consistency constraint $T^{-1}(\mathcal{S}_2 \cap \mathrm{Im}T) \subseteq \mathcal{S}_1$ and allowed by the theory. That is, $T^{-1} \circ T$ is the identity on $\mathcal{S}_1$, and $T \circ T^{-1}$ is the identity on $(\mathcal{S}_2 \cap \mathrm{Im}T)$. This implies that the geometry of $\mathcal{S}_1$ is linearly equivalent to that of $(\mathcal{S}_2 \cap \mathrm{Im}T)$. But not only this. As explained at the end of Section A 1, the physics of $\mathcal{S}_1$ and $(\mathcal{S}_2 \cap \mathrm{Im}T)$ is also equivalent. The invertibility of the map establishes a one-to-one correspondence between states, transformations, measurements and outcome probabilities. For example, if we can implement a particular transformation G in $\mathcal{S}_1$, then we can also implement the corresponding transformation $(T \circ G \circ T^{-1})$ in $(\mathcal{S}_2 \cap \mathrm{Im}T)$. The implementation of $(T \circ G \circ T^{-1})$ consists of mapping a state from $(\mathcal{S}_2 \cap \mathrm{Im}T)$ to $\mathcal{S}_1$, then applying G, and finally mapping the resulting state back via T. A similar argument establishes a correspondence between measurements. Note that the transformation T^{-1} cannot be applied to states in $\mathcal{S}_2$ which are not in $\mathrm{Im}T$. Hence, the implementation of T^{-1} may involve measuring whether the state from $\mathcal{S}_2$ belongs also to $\mathrm{Im}T$ or not.

4. Composite systems

To a setup as the one appearing in Figure 4 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. Two systems A, B constitute a composite system AB if a measurement for A together with a measurement for B uniquely specifies a measurement for AB, independently of the temporal ordering. The fact that subsystems are systems themselves implies that each global state ω_{AB} has well-defined reduced states ω_A, ω_B which do not depend on which transformations and measurements are performed on the other subsystem; this is often referred to as no-signaling. Some bipartite correlations satisfying the no-signaling constraint violate Bell inequalities more than QT does [22]; however, as we will show, these are incompatible with our postulates.

A bipartite system is also a system, so its states can be represented by the probabilities of some fiducial outcomes. But what is the relationship between these and the fiducial outcomes of the subsystems, $x_1, \ldots, x_{k_A}$ and $y_1, \ldots, y_{k_B}$? In order to answer this question, we point out that the fact that $p(x, y)$ does not depend on the ordering of the measurements giving outcomes x, y implies the following lemma.

Lemma 1. The joint probability $p(x, y)$ of any pair of subsystem outcomes x, y is given by

$$p(x, y) = (E_x \otimes E_y) \cdot \omega_{AB} \ , \tag{A2}$$

where

$$\omega_{AB} = \begin{bmatrix} p(x_1, y_1) \\ p(x_1, y_2) \\ \vdots \\ p(x_{k_A}, y_{k_B}) \end{bmatrix} \in \mathbb{R}^{k_A} \otimes \mathbb{R}^{k_B} \ . \tag{A3}$$

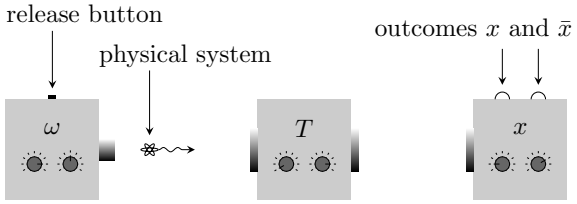

FIG. 4. **General experimental setup.** From left to right there are the preparation, transformation and measurement devices. As soon as the release button is pressed, the preparation device outputs a physical system in the state specified by its knobs. The next device performs the transformation specified 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.

Product states and the set of all these vectors ω_{AB} span the vector space $\mathbb{R}^{k_A} \otimes \mathbb{R}^{k_B}$.

Proof. If the system B is measured first, giving outcome y_j, then the system A is in the state determined by the fiducial probabilities $p(x_i|y_j) = p(x_i, y_j)/p(y_j)$, and the single-system probability rule can be applied $p(x|y_j) = \sum_i E_x^i \, p(x_i|y_j)$. Multiplying by $p(y_j)/p(x)$ and using Bayes' rule gives

$$p(y_j|x) = \sum_i E_x^i \, p(x_i, y_j)/p(x).$$

By using the freedom in the ordering of measurements, we can interpret $p(y_j|x)$ as the state of the system B once the system A has been measured giving outcome x, and the single-system probability rule can be applied again: $p(y_j|x) = \sum_j E_y^j \, p(y_j|x) = \sum_{i,j} E_x^i \, E_y^j \, p(x_i, y_j)/p(x)$. Multiplying both sides of this equality by $p(x)$ gives (A2).

Let us see that the vectors $\omega_{AB} \in \mathcal{S}_{AB}$ span the full tensor product space. In QT, the only states $\omega_{AB} \in \mathcal{S}_{AB}$ which have pure states as marginals $\omega_A \in \mathcal{S}_A, \omega_B \in \mathcal{S}_B$, are product ones $\omega_{AB} = \omega_A \otimes \omega_B$. The same proof technique applies to general probabilistic theories. This implies that $\mathcal{S}_{AB}$ contains all product states, otherwise there would be a state in $\mathcal{S}_A$ or $\mathcal{S}_B$ which is not the marginal of any state in $\mathcal{S}_{AB}$. Next, note that by minimality, $\mathcal{S}_A$ contains k_A linearly independent vectors, and analogously for $\mathcal{S}_B$. The tensor products of these vectors are a set of $k_{AB} = k_A k_B$ linearly independent vectors in $\mathcal{S}_{AB}$, so the set $\mathcal{S}_{AB}$ has full dimension. □

And what about global measurements? The postulate of Tomographic Locality states that the probability for the outcome of any measurement, local or global, is determined by the joint probability $p(x, y)$ of all local measurements. This implies that ω_{AB} in (A3) constitutes a complete representation of a bipartite state, since all outcome probabilities can be calculated from it. Hence, the linear span of $\mathcal{S}_{AB}$ is $\mathbb{R}^{k_A k_B}$, which implies that dimensions follow a multiplicative rule.

$$k_{AB} = k_A k_B \ . \tag{A4}$$

From now on, we use this tensor-product representation given by Eqs. (A2) and (A3) for bipartite states. In this representation, the marginal states are given by $\omega_A = (\mathbb{1}\otimes U)(\omega_{AB})$ and $\omega_B = (U\otimes\mathbb{1})(\omega_{AB})$. For a given pair of states spaces $\mathcal{S}_A, \mathcal{S}_B$ the composite state space $\mathcal{S}_{AB}$ is not unique in general. The only consistency constraints on $\mathcal{S}_{AB}$ are:

1. $\mathcal{S}_{AB}$ must contain the set of separable states, that is the convex hull of $\mathcal{S}_A \otimes \mathcal{S}_B$,

2. all states $\omega \in \mathcal{S}_{AB}$ must give consistent probabilities $(E_x \otimes E_y)(\omega) \in [0,1]$ for all local measurements x, y.

Appendix B: The postulates

In what follows, using the formalism of GPTs, we state our postulates in a formal way and discuss them in more details.

Each of our postulates states that a particular task is in principle possible or impossible. This contrasts with the standard formulation of QT, which has little to do with the possibilities and impossibilities of quantum physics. The operational approach, which we follow here, has been historically successful. Some examples include the impossibility to distinguish between gravitational fields and acceleration (Equivalence principle), the impossibility of any process in which the sole result is the absorption of heat from a reservoir and its complete conversion into work (Second Law of Thermodynamics), and the existence of a fundamental speed limit (Special Relativity).

Postulate 1 (Continuous Reversibility)**.** In any system, the group of transformations $\mathcal{G}$ generated by time-continuous reversible dynamics is transitive on the set of pure states $\mathrm{ext}\mathcal{N}$.

The postulate of Continuous Reversibility was introduced in [4], under the name of "continuity axiom". One of the motivations to assume the reversibility and continuity of time evolution is that the most fundamental theories that we know, classical or quantum, enjoy it. The meaning of continuity here is that, when the system evolves for a very small time, the initial and the final states are almost indistinguishable. This is equivalent to the connectedness of the group of transformations generated by time-continuous dynamics.

Up to present-day experimental accuracy, time evolution seems to be continuous. But it is conceivable that at a small scale it is discrete, and continuity is only an approximation that is valid at sufficiently large scales. In this case, our postulates could be understood as describing the corresponding large-scale effective theory. A very interesting open problem is the classification of theories which satisfy all our postulates except for the continuity part of Postulate 1, that is, when the group of reversible transformations $\mathcal{G}$ is not required to be connected. One theory of this kind is classical probability theory, but it is not known if there are others. In [6] it is shown that, if in addition one assumes the postulate of "Equivalence of Subspaces", which is arguably very strong, the only theories that survive are QT and classical probability theory.

Postulate 2 (Tomographic Locality)**.** The state of a composite system is completely characterized by the correlations of measurements on the individual components: $p(x, y)$ for all local outcomes x, y.

The axiom of Tomographic Locality has a well-defined operational meaning, but additionally, it is mathematically very natural, since it endows state spaces of multipartite systems with the familiar tensor-product structure. The authors of [32] consider ways of relaxing Tomographic Locality.

In order to introduce the third postulate let $\mathcal{S}_{\mathrm{gbit}}$ and $\mathcal{S}^n_{\mathrm{gbit}}$ denote the (yet unknown) state space of a gbit and n gbits, respectively, and k_{gbit} and k^n_{gbit} their corresponding dimensions. According to Tomographic Locality, $k^n_{\mathrm{gbit}} = (k_{\mathrm{gbit}})^n$.

Postulate 3 (Existence of an Information Unit)**.** There is a type of system, *the gbit*, which satisfies the following:

0. Associated to each state space $\mathcal{S}$ there is a number n and a reversible transformation $T_{\mathcal{S}}$ mapping $\mathcal{S}$ to an n-gbit state space $\mathcal{S}^n_{\mathrm{gbit}}$ or subspace.

1. (The state tomography is possible). The state space of a gbit has finite dimension k_{gbit}.

2. (All effects are observable). All effects on a gbit correspond to measurement outcomes.

3. (Gbits interact). The group of transformations generated by time-continuous reversible dynamics of two gbits $\mathcal{G}^2_{\mathrm{gbit}}$ contains an element which is not product ($G_{AB} \neq G_A \otimes G_B$).

4. (Weak Information Causality). If there are four gbit states $\omega_{a,a'} \in \mathcal{S}_{\mathrm{gbit}}$ (with $a, a' \in \{0,1\}$) and an effect E such that $E(\omega_{a,a'}) = \delta_{a,0}$, then any effect E' satisfies $E'(\omega_{a,0}) = E'(\omega_{a,1})$.

Any state of a quantum system can be encoded with arbitrary precision in a sufficient number of classical bits. For instance, this can be achieved by writing its density matrix in a bit string. However, if we are given a physical system in an unknown state, there is no way we can obtain this bit string, unless we are given a large number of copies of the system. By measuring a single copy of the system we could encode the outcome in a bit string, but there is no way we can prepare the same state if the only information we have is this bit string. In other words, this encoding is not reversible.

More generally, the classical bit does not constitute a unit of information capable of reversibly encoding the state of any quantum system, although it does if we restrict to classical systems. However, according to QT, the qubit does constitute such a unit of information, and we think that this is a fundamental aspect of QT. Hence, in this work we promote this to postulate. The fact that there exists an information unit at all is equivalent to saying that all types of information are in some sense equivalent.

Our approach can be summarized in the following way: *Information does play a significant role in the foundations of physics, but we do not say what information actually is.* In

this sense, our postulates specify some properties that information must satisfy, but they do not right away specify its physical implementation. That is, they do not postulate that information must be quantum—instead, this fact is *derived as a consequence* of the properties that information should satisfy.

Let us now discuss the content of Postulate 3. way.

State tomography is possible. In any quantum system, the dimension k and the number of perfectly distinguishable states c are related through the equation $k = c^2$. However, for arbitrary state spaces, the only constraint between the positive integers k and c is $k \geq c$. Hence, although not very natural, it is possible that systems with only two perfectly distinguishable states (like, for instance, gbits) have infinite dimension. However, since for any finite $k_{\rm gbit} > 3$ interaction between gbits is impossible, we are inclined to think that $k_{\rm gbit} = \infty$ is also incompatible with Postulate 3.3. Consequently, we conjecture that Postulate 3.1 is redundant, but, since we cannot prove this fact, we keep the postulate.

Independently of the above discussion, the finiteness of $k_{\rm gbit}$ is necessary if we want state tomography to be possible. The fact that in QT state tomography of infinite-dimensional systems is possible is due to the fact that these systems also have an infinite number of perfectly distinguishable states, and with a bound on the energy, one can effectively consider the system to be finite-dimensional. But this does not work if the infinite-dimensional system has a finite number of distinguishable states, like a gbit.

All effects are observable. A priori, given any state space of a physical system, all effects (i.e., linear functionals that yield valid probabilities between 0 and 1 on all states) describe conceivable outcome probabilities of measurements. However, one might imagine that there are additional physical restrictions, similar to the superselection rules, that somehow render some of the effects impossible to appear in actual measurements. Our postulate says that we do not consider this more complicated situation: we assume that—at least in principle—every effect can appear as the outcome of some measurement.

There is also a reformulation of this postulate which can be used instead, without affecting the conclusions of our work. Instead of all the effects, only effects E for which there are two states $\omega_0, \omega_1 \in \mathcal{S}_{\rm gbit}$ such that $E(\omega_0) = 0$ and $E(\omega_1) = 1$ need to be observable. This second statement is logically equivalent to Chiribella-D'Ariano-Perinotti's information-theoretic postulate named "Perfect Distinguishability" (see [8]), phrased as "every state that is not completely mixed can be perfectly distinguished from some other state", where the operational definition of "completely mixed state" is logically equivalent to not being in the boundary of the state space. Our formulation of Postulate 3.2 is motivated by simplicity.

Gbits interact. Interaction is necessary for the creation of entanglement, and consequently, for the violation of Bell inequalities. But even more, without interaction, classical computation is impossible, since single-gbit gates cannot be universal. More generally, the emergence of structure and complex systems seems impossible in a world without interaction. For these reasons we find it very natural to postulate that gbits interact.

We claim that one can explore what lies beyond QT by relaxing some of our postulates. For example, the family of theories which are compatible with all our postulates except for Postulate 3.3 (i.e., "gbits can interact") is given in [33]. Obviously, in all such theories except for QT there is no entanglement.

Weak Information Causality. Information Causality [10] is a limit on how much complementary information can be simultaneously encoded in a system. On its own, Information Causality suffices to derive Tsirelson's bound [34] and other constraints on quantum correlations [35], but not all of them [11]. Here we show that the full structure of quantum correlations can be derived from Information Causality together with our other postulates.

Our version of Information Causality is called weak Information Causality, and has two differences with respect to the original version. First, the communication task associated to the original formulation is a teleportation analog of ours, where Alice sends Bob classical information (instead of a possibly non-classical gbit) in a context where Bell-violating correlations are shared between both parties. Second, the tradeoff between Bob's knowledge on a and a' that we impose (based on the guessing probability) is weaker than the original one (based on the Shannon mutual information). This has the advantage of having a direct operational meaning, which is missing for the Shannon mutual information in the Information Causality context.

Appendix C: Argumentation

Let $\mathcal{S}$ be the state space of any system allowed by the theory, and k its corresponding dimension. Let $T_{\mathcal{S}} : \mathbb{R}^k \to \mathbb{R}^{k^n_{\rm gbit}}$ be the reversible transformation mapping $\mathcal{S}$ to $(\mathcal{S}^n_{\rm gbit} \cap {\rm Im}T)$ which exists according to Postulate 3.0. As mentioned in Appendix A 3, the physics of $\mathcal{S}$ is equivalent to that of n gbits when restricted to the subspace ${\rm Im}T$, in the sense that there is a perfect correspondence between states, dynamics, measurements, and outcome probabilities. As a consequence, we only need to characterize the state spaces $\mathcal{S}^n_{\rm gbit}$. Once this is done, we will know all state spaces allowed by our postulates.

Our strategy is to show that, the only possible state space $\mathcal{S}^n_{\rm gbit}$ compatible with our postulates is the set of 2^n-dimensional, unit-trace, positive semidefinite, complex matrices ρ; with associated set of effects $\rho \mapsto {\rm tr}M\rho$, where M is any 2^n-dimensional, complex matrix satisfying $0 \leq M \leq \mathbb{1}$; and group of reversible transformations $\rho \mapsto U\rho U^\dagger$, for all $U \in {\rm SU}(2^n)$. In other words, gbits are quantum two-level systems (or qubits). Following the previous paragraph, this implies that all physical systems allowed by our postulates have state spaces being subspaces of $\mathcal{S}^n_{\rm gbit}$, or in other words, can be described within the quantum formalism.

1. A single gbit

Lemma 2. Postulates 3.2 and 3.4 imply that there are no mixed states in the boundary of $\mathcal{N}_{\rm gbit}$.

Proof. Suppose the mixed state $\omega_{\rm mix} = q\omega_1 + (1-q)\omega_2$ is in the boundary of $\mathcal{S}_{\rm gbit}$. Then, there exists an effect E with $E(\omega_{\rm mix}) = 1$ and $E(\omega') = 0$ for some other state $\omega' \in \mathcal{S}_{\rm gbit}$. According to Postulate 3.2 this effect is in principle measurable. Moreover, the linearity of E together with the property $E(\mathcal{S}_{\rm gbit}) \in [0,1]$ imply that $E(\omega_1) = E(\omega_2) = 1$. Therefore, we can encode $a = 0$ in ω_1 or ω_2, and $a = 1$ in ω'. Additionally, we can encode $a' = 0$ in ω_1 and $a' = 1$ in ω_2. Since $\omega_1 \neq \omega_2$, there is an effect E' for which $E'(\omega_1) \neq E'(\omega_2)$. By relabeling $\omega_{0,0} = \omega_1$, $\omega_{0,1} = \omega_2$ and $\omega_{1,0} = \omega_{1,1} = \omega'$ we obtain a contradiction with Postulate 3.4. □

Figure 2 contains a pictorial representation of the above proof.

Lemma 3. Continuous Reversibility together with the fact that $\mathcal{N}_{\rm gbit}$ has no mixed states in its boundary imply that $\mathcal{N}_{\rm gbit}$ is a solid ellipsoid.

Proof. Using the Haar measure on the compact connected Lie group $\mathcal{G}_{\rm gbit}$, we can define a positive matrix

$$W^2 = \int_{\mathcal{G}_{\rm gbit}} dG\; G^{\intercal} G\,, \tag{C1}$$

and W as its unique positive square root. Note that $W^{\intercal} = W$ and $W^2 G^{-1} = G^{\intercal} W^2$ for all $G \in \mathcal{G}_{\rm gbit}$. According to Continuous Reversibility, for any pair of pure states $\omega_1, \omega_2 \in \mathcal{S}_{\rm gbit}$ we have $\omega_2 = G\omega_1$ for some $G \in \mathcal{G}_{\rm gbit}$, and hence

$$\begin{aligned} |W\omega_2| &= \sqrt{\omega_2 \cdot W^2 \omega_2} \\ &= \sqrt{\omega_1 \cdot G^{\intercal} W^2 G \omega_1} \\ &= \sqrt{\omega_1 \cdot W^2 \omega_1} \\ &= |W\omega_1|, \end{aligned}$$

where the notation $\omega_1 \cdot \omega_2$ is used to denote the Euclidean inner product, while, accordingly, $|\omega_1| = \sqrt{\omega_1 \cdot \omega_1}$ stands for the Euclidean norm.

This allows us to define the constant $r = |W\omega|$, where $\omega \in \mathcal{S}_{\rm gbit}$ is a pure state. Note that r is independent of the chosen pure state ω. The set $\mathcal{E} = \{x \in \mathbb{R}^{k_{\rm gbit}} | r = |Wx|\}$ is an ellipsoid, and the intersection of $\mathcal{E}$ and the normalization hyperplane $\mathcal{F} = \{x \in \mathbb{R}^{k_{\rm gbit}} | U \cdot x = 1\}$ is also an ellipsoid. The pure states of $\mathcal{N}_{\rm gbit}$ are contained in the intersection $\mathcal{E} \cap \mathcal{F}$, and since there are no mixed states in the boundary of $\mathcal{N}_{\rm gbit}$, the set of pure states $\mathrm{ext}\mathcal{N}_{\rm gbit}$ must be $\mathcal{E} \cap \mathcal{F}$, which is a $(k_{\rm gbit} - 1)$-dimensional ellipsoid. □

Lemma 4 (Bloch-vector representation). Postulates 1, 3.2 and 3.4 imply the existence of a representation where the state space of a gbit is

$$\mathcal{S}_{\rm gbit} = \left\{ u \begin{bmatrix} 1 \\ \hat{\omega} \end{bmatrix} \,|\, u \in [0,1],\ \hat{\omega} \in \mathbb{R}^d,\ |\hat{\omega}| \leq 1 \right\}, \tag{C2}$$

the normalization effect is $U = [1, \hat{\mathbf{0}}]$, and the group of transformations generated by time-continuous dynamics is

$$\mathcal{G}_{\rm gbit} = \left\{ \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{G} \end{bmatrix} \,|\, \hat{G} \in \hat{\mathcal{G}}_{\rm gbit} \right\}, \tag{C3}$$

where $\hat{\mathcal{G}}_{\rm gbit}$ is a connected subgroup of $\mathrm{SO}(d)$ which is transitive in the unit sphere of $\mathbb{R}^d$, and $d = k_{\rm gbit} - 1 \geq 2$.

Proof. First, we follow the reparametrization procedure explained at the end of Section A 1. In this case, the invertible transformation is $L = \sqrt{2}\, r^{-1} W$, where r and W are defined in the proof of Lemma 3. All the matrices of the reparametrized group $\tilde{\mathcal{G}}_{\rm gbit} = L \circ \mathcal{G}_{\rm gbit} \circ L^{-1}$ are orthogonal. To see this, note that

$$\begin{aligned} \tilde{G}^{\intercal}\tilde{G} &= (WGW^{-1})^{\intercal}(WGW^{-1}) \\ &= \int_{\mathcal{G}_{\rm gbit}} dH\; W^{-1} G^{\intercal} H^{\intercal} H G W^{-1} \\ &= W^{-1} W^2 W^{-1} \\ &= \mathbb{1}\,, \end{aligned}$$

where we have used the fact that W is symmetric. From now on, when referring to the state space, effects and transformations of a gbit, we mean the reparametrized ones,

$$\begin{aligned} \mathcal{S}_{\rm gbit} &\to L(\mathcal{S}_{\rm gbit})\,, \\ U^{\intercal} &\to U^{\intercal} L^{-1}\,, \\ \mathcal{G}_{\rm gbit} &\to L\, \mathcal{G}_{\rm gbit} L^{-1}\,, \end{aligned} \tag{C4}$$

and we omit the tilde.

The orthogonality of the transformations in $\mathcal{G}_{\rm gbit}$ imply that a left eigenvector $U^{\intercal} G = G$ is also a right eigenvector $GU = U$. Hence, the matrix group $\mathcal{G}_{\rm gbit}$ contains a trivial one-dimensional representation spanned by U, and another representation denoted $\hat{\mathcal{G}}_{\rm gbit}$. So, for any $G \in \mathcal{G}_{\rm gbit}$ there is $\hat{G} \in \hat{\mathcal{G}}_{\rm gbit}$ such that

$$G = \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & \hat{G} \end{bmatrix}\,. \tag{C5}$$

In our notation, symbols with a hat "ˆ" are associated to the non-trivial representation of $\mathcal{G}_{\rm gbit}$. In this basis, the normalization effect is $U = [1, \hat{\mathbf{0}}]$, and the pure states are $\omega = [1, \hat{\omega}]$ with $|\hat{\omega}| = 1$, where the latter is a consequence of the fact that, according to the definition of L, pure states have Euclidean norm $|\omega| = \sqrt{2}$. □

From now on, when dealing with a gbit, we adopt the representation given in Lemma 4. Note that, in this representation, pure states are those with unit normalization $u = 1$, and unit-length Bloch vector $|\hat{\omega}| = 1$. Each effect is characterized by a vector $E = [e, \hat{E}]$ such that

$$E(\omega) = E \cdot \omega = u(e + \hat{E} \cdot \hat{\omega})\,. \tag{C6}$$

The consistency constraint $E(\mathcal{S}_{\rm gbit}) \subseteq [0,1]$ is equivalent to $|\hat{E}| \leq e$ and $e + |\hat{E}| \leq 1$. An effect E for which there are

two states $\omega_0, \omega_1 \in \mathcal{S}_{\rm gbit}$ such that $E(\omega_0) = 0$ and $E(\omega_1) = 1$, satisfies $e = |\hat{E}| = 1/2$. Such effects are in one-to-one correspondence with pure states $\omega \in {\rm ext}\mathcal{N}_{\rm gbit}$ via the map $E = \omega/2$.

2. Two gbits

Tomographic Locality and Lemma 4 imply that two-gbit states can be represented as

$$\omega_{AB} = u \begin{bmatrix} 1 \\ \alpha \\ \beta \\ \gamma \end{bmatrix} \in \mathcal{S}^2_{\rm gbit} \,, \tag{C7}$$

where $\alpha = \hat{\omega}_A \in \mathbb{R}^d$, $\beta = \hat{\omega}_B \in \mathbb{R}^d$, and $\gamma \in \mathbb{R}^d \otimes \mathbb{R}^d$ is called the "correlation matrix". Note that the ordering of the components in (C7) is different from the one in (A3). At this stage, we know that $|\alpha|, |\beta| \leq 1$, but we do not know much about the full structure of $\mathcal{S}^2_{\rm gbit}$, nor its associated group $\mathcal{G}^2_{\rm gbit}$. However, these two objects are very much related. Indeed, the postulate of Continuous Reversibility implies that the set of pure states for two gbits is ${\rm ext}\mathcal{N}^2_{\rm gbit} = \mathcal{G}^2_{\rm gbit}(\omega \otimes \omega)$, where $\omega \in {\rm ext}\mathcal{N}_{\rm gbit}$ is any pure state. In order to see this, recall that product states belong to $\mathcal{S}^2_{\rm gbit}$, and that the product of two locally pure states is a globally pure state. This connection between $\mathcal{S}^2_{\rm gbit}$ and $\mathcal{G}^2_{\rm gbit}$ implies that the consistency constraints for $\mathcal{S}^2_{\rm gbit}$ mentioned at the end of Section A 4, translate to constraints for $\mathcal{G}^2_{\rm gbit}$. These constraints are the premise of the following lemma.

Lemma 5. Let $\hat{\mathcal{G}}_{\rm gbit}$ be a connected subgroup of ${\rm SO}(d)$ which is transitive in the unit sphere of $\mathbb{R}^d$, where $d \geq 2$. Let $\mathcal{G}^2_{\rm gbit}$ be a connected group of real $(d+1)^2 \times (d+1)^2$ matrices which satisfies the following:

1. $(\mathcal{G}_{\rm gbit} \otimes \mathcal{G}_{\rm gbit}) \leq \mathcal{G}^2_{\rm gbit}$,
2. $(E \otimes E) \cdot G(\omega \otimes \omega) \in [0,1]$ for all $G \in \mathcal{G}^2_{\rm gbit}$,

where $\omega = [1, \hat{\omega}]$, $|\hat{\omega}| = 1$ and $E = \omega/2$. If $d \neq 3$, then the group $\mathcal{G}^2_{\rm gbit}$ must be a subgroup of $\mathcal{H}_d \otimes \mathcal{H}_d$, where

$$\mathcal{H}_d = \left\{ \begin{bmatrix} 1 & \mathbf{0} \\ \mathbf{0} & Q \end{bmatrix} \mid Q \in {\rm SO}(d) \right\} \,. \tag{C8}$$

Proof. See Ref. [33]. □

Lemma 5 shows that, except for $d = 3$, the joint state space $\mathcal{S}^2_{\rm gbit}$ only contains separable states, and its associated group $\mathcal{G}^2_{\rm gbit}$ only contains non-interacting transformations, which is in contradiction with Postulate 3.3. Hence, the only possibility is $d = 3$, and this is getting quite close to QT. It turns out that the only subgroup of ${\rm SO}(3)$ which is transitive on the unit sphere of $\mathbb{R}^3$ is ${\rm SO}(3)$ itself. Hence, from now on, we assume $d = 3$ and $\hat{\mathcal{G}}_{\rm gbit} = {\rm SO}(3)$.

3. Emergence of quantum theory

Since QT satisfies our postulates, it must fit the structure that we have found up to this stage. Indeed, the state of a qubit can be represented by the three-dimensional Bloch vector $\hat{\omega}$ (and the normalization parameter u if we consider general states). That is, the state space $\mathcal{S}_{\rm gbit}$ of a gbit and that of a qubit, $\mathcal{S}_{\rm qubit}$, are equivalent in the sense of Subsection A 1: both are in one-to-one correspondence via an invertible linear map L. It is given by

$$\begin{aligned} L : \mathcal{S}_{\rm gbit} &\to \mathcal{S}_{\rm qubit} \\ \begin{bmatrix} 1 \\ \hat{\omega} \end{bmatrix} &\mapsto \frac{1}{2} \left(1 \cdot \mathbb{1} + \hat{\omega} \cdot \vec{\sigma} \right) , \end{aligned}$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_y)$ is a vector with the Pauli matrices as entries. This maps Bloch vectors onto density matrices, $\mathcal{S}_{\rm qubit} = \{\rho \in \mathbb{C}^{2\times 2} \mid \rho \geq 0, {\rm tr}(\rho) = 1\}$. The group of reversible transformations for one gbit is $\hat{\mathcal{G}}_{\rm gbit} = {\rm SO}(3)$, which is equivalent to the adjoint representation of ${\rm SU}(2)$:

$$\mathcal{G}_{\rm qubit} = L\mathcal{G}_{\rm gbit}L^{-1} = \left\{ \rho \mapsto U\rho U^\dagger \mid U \in SU(2) \right\} .$$

The group of reversible transformations for n qubits, denoted $\mathcal{G}^n_{\rm qubit}$, is the adjoint action of ${\rm SU}(2^n)$:

$$\mathcal{G}^n_{\rm qubit} = \left\{ \rho \mapsto U\rho U^\dagger \mid U \in SU(2^n) \right\} .$$

As shown in [33], this is the *only* possible choice of any n-gbit dynamics which satisfies our postulates—up to an equivalence transformation, which is L for a single gbit (mapping Bloch vectors to density matrices), and, correspondingly, $L^{\otimes n} = L \otimes \ldots \otimes L$ for n gbits, mapping the corresponding state vectors to density matrices of size 2^n:

Lemma 6. Let $\hat{\mathcal{G}}_{\rm gbit} = {\rm SO}(3)$ and let $\mathcal{G}^n_{\rm gbit}$ be a connected group of $(3+1)^n \times (3+1)^n$ real matrices which satisfies the following:

1. $(\mathcal{G}_{\rm gbit} \otimes \cdots \otimes \mathcal{G}_{\rm gbit}) \leq \mathcal{G}^n_{\rm gbit}$,
2. $(E \otimes \cdots \otimes E) \cdot G(\omega \otimes \cdots \otimes \omega) \in [0,1]$ for all $G \in \mathcal{G}^n_{\rm gbit}$,

where $\omega = [1, \hat{\omega}]$, $|\hat{\omega}| = 1$ and $E = \omega/2$. The only possible groups $\mathcal{G}^n_{\rm gbit}$ are:

1. $\mathcal{G}_{\rm gbit} \otimes \cdots \otimes \mathcal{G}_{\rm gbit}$
2. $\left(L^{-1}\right)^{\otimes n} \mathcal{G}^n_{\rm qubit} L^{\otimes n}$.

Proof. See Ref. [33]. □

In summary, we have shown that according to our postulates, an n-gbit system is equivalent to a quantum system with Hilbert space dimension 2^n. Now, Postulate 3.0 says that any state space is reversibly encodable in a multi-qubit system. Hence, the states, transformations and measurements of any system allowed by our postulates can be described within the formalism of QT.