Editors' Suggestion

Entanglement and Superposition Are Equivalent Concepts in Any Physical Theory

Guillaume Aubrun,[1,*] Ludovico Lami,[2,†] Carlos Palazuelos,[3,4,‡] and Martin Plávala[5,§]

[1] *Institut Camille Jordan, Université Claude Bernard Lyon 1, 43 boulevard du 11 novembre 1918, 69622 Villeurbanne CEDEX, France*

[2] *Institute of Theoretical Physics and IQST, Universität Ulm, Albert-Einstein-Allee, 11D-89069 Ulm, Germany*

[3] *Departamento de Análisis Matemático y Matemática Aplicada, Universidad Complutense de Madrid, Plaza de Ciencias s/n, 28040 Madrid, Spain,*

[4] *Instituto de Ciencias Matemáticas, c/ Nicolás Cabrera, 13-15, 28049 Madrid, Spain*

[5] *Naturwissenschaftlich-Technische Fakultät, Universität Siegen, 57068 Siegen, Germany*

(Received 20 September 2021; revised 7 December 2021; accepted 9 March 2022; published 22 April 2022)

We prove that given any two general probabilistic theories (GPTs) the following are equivalent: (i) each theory is nonclassical, meaning that neither of their state spaces is a simplex; (ii) each theory satisfies a strong notion of incompatibility equivalent to the existence of "superpositions"; and (iii) the two theories are entangleable, in the sense that their composite exhibits either entangled states or entangled measurements. Intuitively, in the post-quantum GPT setting, a superposition is a set of two binary ensembles of states that are unambiguously distinguishable if the ensemble is revealed before the measurement has occurred, but not if it is revealed after. This notion is important because we show that, just like in quantum theory, superposition in the form of strong incompatibility is sufficient to realize the Bennett-Brassard 1984 protocol for secret key distribution.

DOI: 10.1103/PhysRevLett.128.160402

Introduction.—When one looks back at the magnificent conceptual revolution that quantum mechanics sparked almost a century ago, two discoveries stand out as marking a decisive departure from classical physics, namely, the superposition principle and the existence of entanglement. The former can be exploited to design powerful cryptographic protocols [1,2], while the latter implies, via Bell's theorem [3,4], that the correlations exhibited by separate systems cannot be explained by means of local hidden variable models. These consequences of superposition and entanglement are predicted by the formalism of quantum mechanics but can also be understood operationally, as simple statements concerning the frequencies of certain measurement outcomes. They can thus be regarded as theory independent: any future "ultimate" theory of nature, which may overcome quantum mechanics, must encompass them and explain those experiments.

What is theory dependent is the *connection* between these two notions. Namely, it is only within the formalism of quantum theory that we can understand entanglement as the superposition principle applied to different product vectors of a tensor product Hilbert space [5,6]. The fact that the connection between two fundamental and experimentally verified phenomena can only be understood by means of the mathematical formalism pertaining to a specific framework is conceptually unsatisfying. What is more, this makes our understanding more dependent on the current theoretical paradigm—which is, most likely, incomplete. And indeed, recently there have been several attempts to investigate the interplay between these two notions in an *a priori* fashion [7–12].

In this Letter we introduce the concept of *strong incompatibility*, showing that it leads naturally to the sought-after theory-independent connection between superposition and entanglement. This promotes such a connection from a mere accident of the mathematics underpinning quantum mechanics to a logical necessity. This connection is illustrated in Fig. 1, which depicts how strong incompatibility as given by Theorem 2 connects nonclassicality and superposition with entangleability, and at the same time allows us to construct a version of the Bennett-Brassard 1984 (BB84) cryptographic protocol [1] in any nonclassical theory.

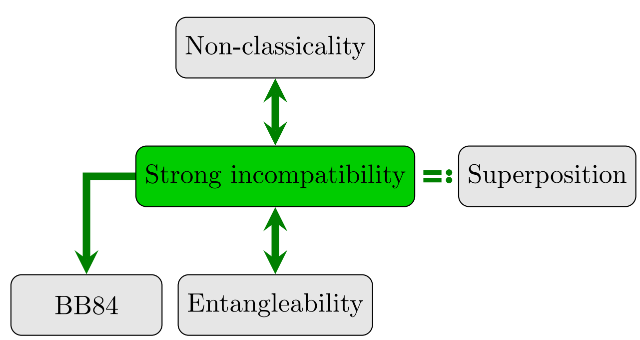

FIG. 1. Connections between the notions of nonclassicality and superposition and entangleability via that of strong incompatibility, which is also key to implementing the BB84 protocol.

To pursue our program, we need a framework capable of encompassing all physical theories obeying minimal operational requirements, beyond standard quantum theory. The formalism of general probabilistic theories (GPTs) accomplishes precisely that [13–18]. A brief introduction can be found below or in Refs. [19–22]. Before explaining our result, we need to answer a few questions.

(I) *What does it mean that a certain theory is nonclassical?* The answer we shall adopt is that its state space should not be described by a classical probability theory: that is, there should not be a finite set of "elementary states" that are both (a) perfectly distinguishable by a measurement; and (b) such that any other state can be written as a statistical mixture of them. In the geometric language of GPTs, this amounts to saying that the state space is not shaped as a simplex, the multidimensional generalization of the triangle and tetrahedron.

(II) *How can we define superposition without reference to quantum theory?* The difficulty here lies in the fact that without a Hilbert space, not available in a generic GPT, there is no notion of linear combination of state vectors. However, we can try to follow a different route: the operational notion of superposition stems from the comparison between the two ensembles of single-qubit quantum states $\{|0\rangle, |1\rangle\}$ and $\{|+\rangle, |-\rangle\}$, where $|\pm\rangle = (1/\sqrt{2})(|0\rangle \pm |1\rangle)$. These ensembles, when mixed with equal weights, give rise to the same density matrix; further, identifying the state unambiguously via a measurement is possible or impossible depending on whether the ensemble is revealed before or after the measurement is carried out. We will use such ensembles, termed strongly incompatible, as a way to identify the existence of superpositions in an arbitrary GPT. Theorem 2 below shows that this is a meaningful ansatz, as superpositions exist if and only if the theory is nonclassical.

(III) *What does it mean that two theories exhibit entanglement when combined*? Let us distinguish between entanglement at the level of states and of measurements. The former means that there are states on the bipartite system that cannot be written as a statistical mixture of uncorrelated (product) states. The latter, accordingly, means that there are bipartite effects that are not a positive linear combination of product effects.

General probabilistic theories.—In the most general sense, a physical theory is simply a set of rules that allow to deduce a probabilistic prediction of the outcome of an experiment given the detailed description of its preparation. From this abstract description one can deduce, via Ludwig's embedding theorem [14,15,20], the mathematical formalism of general probabilistic theories that we will now describe [20–22].

The fundamental object needed to model an arbitrary physical system is its *state space*; mathematically, this will be represented by a generic convex and compact subset Ω of some finite-dimensional real vector space [23]. Physically, a state $\omega \in \Omega$ should be thought of as a description of a preparation procedure for the system under examination. The convexity of Ω reflects the fact that preparation procedures can be mixed stochastically: the ensemble $\{p_i, \omega_i\}$, which corresponds to the physical procedure of drawing a random variable I and preparing the system in the state ω_i, is represented within the formalism by the convex mixture $\sum_i p_i \omega_i$.

It is useful to include into the picture not only normalized but also un-normalized states. To do so, we imagine an augmented vector space V where we introduce a proper cone C, i.e., a set $C \subset V$ that is closed under positive scalar multiplication, and moreover: (i) convex; (ii) salient, meaning that $C \cap (-C) = \{0\}$; (iii) generating, in the sense that $C - C = V$; and (iv) topologically closed (Fig. 2). The state space Ω is then recovered as the section of C identified by the equation $u = 1$, where $u \in V^*$ is a "normalizing" functional, called the *order unit*, belonging to the dual vector space V^* and (v) strictly positive on C, i.e., such that $u(x) > 0$ for all $x \in C$ with $x \neq 0$. We can summarize the above discussion by giving an abstract definition of a GPT as any triple (V, C, u), where V is a real finite-dimensional vector space, $C \subset V$ is a proper cone inside it, and $u \in V^*$ is a strictly positive functional on C.

This makes V an ordered vector space: for any two $x, y \in V$, we define the ordering by stipulating that $x \leq y$ if $y - x \in C$. Notably, this ordering is not total, i.e., it is possible that neither $x \leq y$ nor $y \leq x$. The dual space V^* inherits an ordering from V: for $f, g \in V^*$, we write $f \leq g$ if $f(\omega) \leq g(\omega)$ for all $\omega \in \Omega$ [equivalently, $f(x) \leq g(x)$ for all $x \in C$]. The cone of positive functionals in V^*, called the *dual cone* to C, is denoted with C^*. Remarkably, for proper cones C we have the identity $C^{**} = C$.

To complete our picture we need to discuss measurements. A physical measurement together with one of its possible outcomes will be represented mathematically by an *effect*. This is just a linear functional $e \in V^*$; for every state $\omega \in \Omega$, the value $e(\omega)$ is interpreted as the probability that the corresponding outcome occurs when that state is measured, and thus $0 \leq e(\omega) \leq 1$. Employing the above

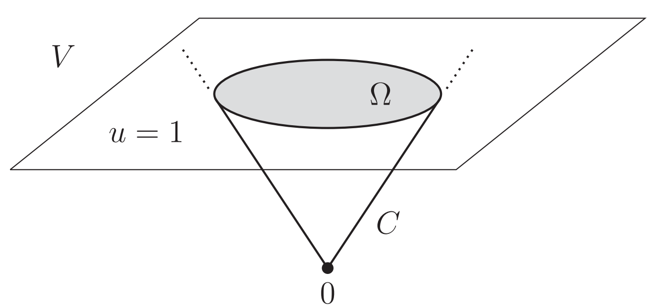

FIG. 2. The basic ingredients of a GPT are a real finite-dimensional vector space V and a cone C. The order unit functional u defines a hyperplane $u^{-1}(1)$, whose intersection with C identifies the state space Ω.

notion of ordering on V^*, we see that $0 \le e \le u$. A fully fledged measurement will then be a (finite) collection of effects $(e_i)_{i\in I}$, where $e_i \in V^*$ with $e_i \ge 0$ for all i, and moreover $\sum_{i\in I} e_i = u$, so that the outcome probabilities add up to 1.

Is this mathematical description of measurements complete? Namely, given a collection of effects summing to u, can it be physically implemented as a measurement procedure? If so, the system is said to satisfy the no-restriction hypothesis [24,25]. We deem this assumption natural, for the following reasons: classical theories (example 1) and quantum mechanics satisfy it; assuming additional restrictions is unnecessary and thus should be avoided whenever possible; the no-restriction hypothesis does not limit the mathematical validity of our results, as in theories with restrictions one only needs to check whether the needed effects are available; finally, dropping this assumption trivializes the problem from the mathematical standpoint (see below). Thus, throughout this Letter we will always include the no-restriction hypothesis in our theoretical framework.

To make the GPT formalism we have sketched more concrete, let us discuss an important example of it.

Example 1: Classical theories as GPTs.—In classical theory the cone C is generated by a set of linearly independent states. It then follows that the state space is a simplex and that every state is given as a unique convex combination of the generating states.

Strong incompatibility.—We consider a strengthened version of the well-known notion of incompatibility and prove that it is in fact fully equivalent to non-classicality (Fig. 1). The incompatibility we have in mind here is both at the level of states and at the level of effects. While the latter notion is the more commonly studied, its operational interpretation resting on the fact that two measurements are compatible if and only if they can be implemented jointly [26], one also encounters the concept of incompatibility of states when investigating whether a given assemblage has a local hidden state model [27]. Moreover, within the context of GPTs the cones generated by states and effects are dual to each other, so it is natural that the two versions of incompatibility be treated on equal footing. Given a vector space V ordered by a cone C, two finite families of vectors $x_i \in C$ and $y_j \in C$ are said to be compatible if one can find $z_{ij} \in C$, such that $\sum_j z_{ij} = x_i$ and $\sum_i z_{ij} = y_j$ for all i, j; they are said to be incompatible otherwise. Clearly, a necessary but in general not sufficient condition for compatibility is that $\sum_i x_i = \sum_j y_j$. If (V, C, u) forms a GPT, we can try to find incompatible vectors either in the primal space V, C or in the dual space V^*, C^*.

The connection between incompatibility and nonclassicality of GPTs has been explored thoroughly [9,28–36]. For instance, it is known that a GPT is nonclassical if and only if it admits two incompatible binary measurements [34,36]. Here we establish a modified and stronger version of this fact:

Theorem 2.—A proper cone C is nonclassical if and only if there are vectors $0 \ne x_0, x_1, x_+, x_- \in C$ and functionals $f_0, f_1, f_+, f_- \in C^*$ such that (i) $x_0 + x_1 = x_+ + x_-$ and $f_0 + f_1 = f_+ + f_-$; (ii) $f_0(x_1) = f_1(x_0) = f_+(x_-) = f_-(x_+) = 0$; and (iii) $f_i + f_j$ is strictly positive, for all $i \in \{0, 1\}$, $j \in \{+, -\}$.

The proof of Theorem 2 is in the Supplemental Material [37]. At first sight it may not be clear what Theorem 2 has to do with the notion of incompatibility. However, the two families of vectors x_0, x_1; x_+, x_- constructed there are in fact incompatible. Indeed, assume that a decomposition $(z_{ij})_{ij} \in C$ holds, so that $\sum_j z_{ij} = x_i$ and $\sum_i z_{ij} = x_j$. Then $0 = f_1(x_0) = f_1(z_{0+} + z_{0-}) \ge f_1(z_{0+}) \ge 0$ and analogously $0 = f_-(x_+) = f_-(z_{0+} + z_{1+}) \ge f_-(z_{0+}) \ge 0$, so that $f_1(z_{0+}) = f_-(z_{0+}) = 0$. Since $f_1 + f_-$ must be strictly positive and $(f_1 + f_-)(z_{0+}) = 0$, it holds that $z_{0+} = 0$. Repeating this reasoning we reach the absurd conclusion that $z_{ij} \equiv 0$ for all i, j; hence, the vectors x_0, x_1; x_+, x_- were incompatible.

As we will show in the next section, Theorem 2 solidifies our idea of defining superposition as described in the Introduction [see (II)]. Indeed, one can draw a direct parallel between the two families of vectors x_0, x_1; x_+, x_- and the quantum states $|0\rangle, |1\rangle; |+\rangle, |-\rangle$ representing states of a qubit, where $|\pm\rangle = (1/\sqrt{2})(|0\rangle \pm |1\rangle)$. The corresponding effects f_0, f_1; f_+, f_- are then simply the projections onto the subspaces generated by the respective vectors. In this sense, Theorem 2 implies that any nonclassical state space exhibits an operational form of discord [38,39].

BB84 protocol in GPTs.—As the main application of the theory developed here, we design a version of the BB84 protocol [1] for secret key distribution over a public channel that works in any nonclassical GPT. The idea comes from the aforementioned parallel between the families of vectors x_0, x_1; x_+, x_- and the quantum states $|0\rangle$, $|1\rangle$; $|+\rangle$, $|-\rangle$. Since the latter are employed in the quantum BB84 protocol, it is natural to ask whether the former can be used in a similar way in any nonclassical GPT (V, C, u). To this end, let $x_0, x_1; x_+, x_-$ and $f_0, f_1; f_+, f_-$ be the vectors given by Theorem 2. Construct states ρ_0, ρ_1; $\sigma_+, \sigma_- \in \Omega$ such that $p_i\rho_i = x_i$ and $q_j\sigma_j = x_j$ for some $p_i, q_j > 0$. Then, $p_0\rho_0 + p_1\rho_1 = q_+\sigma_+ + q_-\sigma_-$. By rescaling if necessary, we can assume that $p_0 + p_1 = q_+ + q_- = 1$ and similarly that $f_0 + f_1 = f_+ + f_- =: \ell \le u$.

Now, Alice tosses a fair coin; if heads, she prepares one of the states ρ_0, ρ_1 (with *a priori* probabilities p_0, p_1); if tails, one of the states σ_+, σ_- (with *a priori* probabilities q_+, q_-). Since $p_0\rho_0 + p_1\rho_1 = q_+\sigma_+ + q_-\sigma_-$, an eavesdropper Eve cannot discern these two scenarios. Unlike in the quantum case, it is not guaranteed that Bob can perfectly discriminate ρ_0, ρ_1 or σ_+, σ_-; however, he will toss a fair coin too, and run an unambiguous state discrimination procedure using the measurements $f_0, f_1, u - \ell$ (if heads) or $f_+, f_-, u - \ell$ (if tails). This introduces an additional error, as the rounds where Bob obtains the outcome $u - \ell$ have to be discarded. Despite that, Alice and Bob can proceed as usual: they make the results of their coin tosses public and remove the rounds for which either the choices of preparation and measurement were different or Bob obtained the outcome $u - \ell$. In the remaining cases the choices of preparation and measurement correspond, and Bob's outcome was not $u - \ell$. By Theorem 2(ii), Bob has thus recovered with no error the key bit i. In this way Alice and Bob obtain a shared key. One of the significant differences with the quantum case is that this key is not automatically secret. In fact, the information revealed to Eve *is* correlated with the key bit. To remedy this, Alice and Bob can run the secret key distillation protocol proposed by Maurer [40] to extract a truly secure key. A detailed description of this step as well as proof that it achieves a nonzero secret key rate can be found in the Supplemental Material [37].

Bipartite systems.—In order to describe entanglement we need to discuss bipartite systems first. Given two systems A, B modeled by GPTs $A = (V_1, C_1, u_1)$ and $B = (V_2, C_2, u_2)$, can we represent also the joint system AB as a GPT $AB = (V_{12}, C_{12}, u_{12})$? In this context, a natural assumption—which we shall adopt throughout the Letter—is the *local tomography principle*. In layman's terms, it states that the composite system should not contain more degrees of freedom than its parts. More formally, we require that the statistics under product measurements determine any state of the bipartite system uniquely. With this assumption, one can prove the familiar tensor product rule [41,42]

$$V_{12} = V_1 \otimes V_2, \qquad u_{12} = u_1 \otimes u_2. \tag{1}$$

Two operationally motivated constraints on the cone C_{12} come from the fact that independent local actions, namely, state preparations and measurements, should be faithfully represented in the bipartite picture as well. More formally, (i) local (tensor product) states should also be valid bipartite states, and (ii) local (tensor product) effects should also be valid effects on the bipartite system. Introducing the *minimal* and the *maximal* tensor product of the cones C_1 and C_2, defined by

$$C_1 \underset{\min}{\otimes} C_2 \coloneqq \mathrm{conv}\{x \otimes y : x \in C_1, y \in C_2\}, \tag{2}$$

$$C_1 \underset{\max}{\otimes} C_2 = \left(C_1^* \underset{\min}{\otimes} C_2^* \right)^*, \tag{3}$$

where conv denotes the convex hull, we can rephrase (i) as $C_1 \otimes_{\min} C_2 \subseteq C_{12}$ and (ii) as $C_1^* \otimes_{\min} C_2^* \subseteq C_{12}^*$. By combining the former relation with the dual of the latter we obtain the twofold bound

$$C_1 \underset{\min}{\otimes} C_2 \subseteq C_{12} \subseteq C_1 \underset{\max}{\otimes} C_2 \tag{4}$$

on the bipartite cone C_{12}. We can now formalize the answer to question (III) in the Introduction: the existence of entanglement at the level of states or measurements is equivalent to one of the two inclusions in Eq. (4) being strict. In turn, this happens if and only if

$$C_1 \underset{\min}{\otimes} C_2 \neq C_1 \underset{\max}{\otimes} C_2, \tag{5}$$

i.e., if the minimal tensor product is a strict subset of the maximal tensor product. When this is the case we call A, B *entangleable*. One interesting aspect of this definition of entangleability is that it is independent of the bipartite cone: whatever C_{12} is, Eq. (5) guarantees that the joint system will exhibit either entangled states or entangled measurements (or both).

Entangleability.—Our result on entangleability is as follows:

Theorem 3.—Two GPTs A, B are entangleable if and only if they are both nonclassical.

A somewhat abstract proof of the above result can be found in Ref. [43]. However, thanks to Theorem 2 introduced here, we can now present a more intuitive proof of Theorem 3 [37]. Indeed, given two nonclassical GPTs $A = (V_1, C_1, u_1)$ and $B = (V_2, C_2, u_2)$, thanks to Theorem 2 we can construct an explicit tensor belonging to $C_1 \otimes_{\max} C_2$ but not to $C_1 \otimes_{\min} C_2$, thus demonstrating Eq. (5). In order to do this, we invoke Theorem 2 for the cone C_1 (C_2) to construct vectors $0 \neq x_0, x_1; x_+, x_- \in C_1$ and functionals $f_0, f_1; f_+, f_- \in C_1^*$ (vectors $y_0, y_1; y_+, y_- \in C_2$ and functionals $g_0, g_1; g_+, g_- \in C_2^*$) satisfying conditions (i)–(iii). We then construct the state $\omega = x_0 \otimes y_+ - x_+ \otimes y_+ + x_+ \otimes y_0 + x_1 \otimes y_1$. It turns out that $\omega \in C_1 \otimes_{\max} C_2 \setminus C_1 \otimes_{\min} C_2$, thus implying Eq. (5), i.e., A, B are entangleable. The proof of this claim consists of two parts: first we show that $\omega \in C_1 \otimes_{\max} C_2$, which follows straight from (i). To demonstrate that $\omega \notin C_1 \otimes_{\min} C_2$ we construct a Bell-like inequality of the Clauser-Horne-Shimony-Holt (CHSH) type [44], using the functionals $f_0, f_1; f_+, f_-$, and prove that it is violated. Since in general $f_0 + f_1 = f_+ + f_- \neq u$, the aforementioned Bell-like inequality is not necessarily a Bell inequality in the underlying GPTs, and the question whether any two nonclassical GPTs violate some Bell inequality is still open. The other implication of Theorem 3 is easy: if either $A = (V_1, C_1, u_1)$ or $B = (V_2, C_2, u_2)$ is classical, then it can be seen directly that $C_1 \otimes_{\min} C_2 = C_1 \otimes_{\max} C_2$, i.e., A and B are not entangleable [45,46].

The above Theorem 3 pinpoints a profound and intrinsic connection between nonclassicality—and thus, superposition—and entanglement: the two notions are not merely linked by a mathematical accident of the quantum mechanical formalism but rather two sides of the same coin. Theorem 3 relies on two main assumptions: first, the no-restriction hypothesis, positing that every mathematically consistent effect is physically realizable; and second, the local tomography principle, which entails that combining two systems does not lead to the appearance of new degrees of freedom. These two assumptions are not only natural but also necessary to avoid the mathematical trivialization of the problem. In fact, by dropping the no-restriction hypothesis it is possible to enforce a minimal tensor product composition rule at the level of states and of measurements at the same time, eliminating entanglement somewhat artificially. On the other hand, without local tomography the dimension of the linear span of C_{12} is larger than that of the span of $C_1 \otimes_{\min} C_2$, directly implying the existence of entangled states [10] (Proposition 2).

Conclusions.—We have showed that superpositions are present in every nonclassical GPT and that even in operational settings they allow us to prove existence of BB84 protocol and entanglement. Our results, bypassing the Hilbert space structure of quantum theory, give a counterexample to possible axiomatizations of it [47]: for example, it is known that existence of purifications [25,48], certain symmetries [49,50] or a strong symmetry condition and spectrality [51,52] are enough to single out quantum(-like) theories among other nonclassical theories. Our results show that existence of superpositions, entanglement, and availability of BB84 protocol do not restrict the set of possible nonclassical theories at all.

Our techniques rely on the novel notion of strong incompatibility, which enabled the construction of a universally entangled tensor and of a generalized version of the BB84 protocol in any nonclassical GPT. It is an open question whether one can derive other properties such as no broadcasting [53] from strong incompatibility. It is also open whether the violations of Bell inequalities and steering exist in any nonclassical GPT; we anticipate that strong incompatibility may play an important role in investigating this question.

G. A. was supported in part by ANR (France) under the Grants No. StoQ (2014-CE25-0003) and No. ESQuisses (ANR-20-CE47-0014-01). L. L. acknowledges financial support from the European Research Council under the Starting Grant GQCOP (Grant No. 637352), from the Foundational Questions Institute under the Grant FQXi-RFP-IPW-1907, and from the Alexander von Humboldt Foundation. C. P. is partially supported by Spanish Ministerio de Ciencia e Innovación through Grants No. PID2020-113523GB-I00 and No. CEX2019-000904-S, funded by MCIN/AEI/10.13039/501100011033 and by the Comunidad de Madrid through Grant No. QUITEMADCMS2018/TCS-4342. M. P. acknowledges support from the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, Projects No. 447948357 and No. 440958198), the Sino-German Center for Research Promotion (Project M-0294), the ERC (Consolidator Grant 683107/TempoQ), and from the Alexander von Humboldt Foundation.

[*] aubrun@math.univ-lyon1.fr
[†] ludovico.lami@gmail.com
[‡] carlospalazuelos@mat.ucm.es
[§] martin.plavala@uni-siegen.de

[1] C. H. Bennett and G. Brassard, *Proceedings of the IEEE International Conference on Computers, Systems and Signal Processing*, Bangalore, India (1984), pp. 175–179.
[2] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).
[3] J. Bell, Physics **1**, 195 (1964).
[4] N. Brunner, D. Cavalcanti, S. Pironio, V. Scarani, and S. Wehner, Rev. Mod. Phys. **86**, 419 (2014).
[5] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[6] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).
[7] J. Oppenheim and S. Wehner, Science **330**, 1072 (2010).
[8] J. G. Richens, J. H. Selby, and S. W. Al-Safi, Phys. Rev. Lett. **119**, 080503 (2017).
[9] A. Jenčová, Phys. Rev. A **98**, 012133 (2018).
[10] G. M. D'Ariano, M. Erba, and P. Perinotti, Phys. Rev. A **101**, 042118 (2020).
[11] G. Aubrun, L. Lami, and C. Palazuelos, arXiv:1910.04745.
[12] G. Aubrun, L. Lami, C. Palazuelos, S. J. Szarek, and A. Winter, Commun. Math. Phys. **375**, 679 (2020).
[13] A. Hartkämper and H. Neumann, *Foundations of Quantum Mechanics and Ordered Linear Spaces: Advanced Study Institute held in Marburg 1973* (Springer, Berlin, Heidelberg, 1974).
[14] G. Ludwig, *An Axiomatic Basis for Quantum Mechanics: Derivation of Hilbert space structure*, Vol. 1 (Springer-Verlag, Berlin, 1985).
[15] G. Ludwig, Z. Phys. **181**, 233 (1964).
[16] G. Ludwig, Commun. Math. Phys. **4**, 331 (1967).
[17] G. Ludwig, Commun. Math. Phys. **9**, 1 (1968).
[18] E. Davies and J. Lewis, Commun. Math. Phys. **17**, 239 (1970).
[19] L. Lami, C. Palazuelos, and A. Winter, Commun. Math. Phys. **361**, 661 (2018).
[20] L. Lami, Non-classical correlations in quantum mechanics and beyond, Ph.D. thesis, Universitat Autònoma de Barcelona, 2017.
[21] M. Müller, SciPost Phys. Lect. Notes **28** (2021).
[22] M. Plávala, arXiv:2103.07469.
[23] The assumption of finite dimension is a technical one. It is possible and in general desirable to drop it [20], Chap. 1.
[24] P. Janotta and R. Lal, Phys. Rev. A **87**, 052131 (2013).
[25] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **81**, 062348 (2010).
[26] T. Heinosaari, T. Miyadera, and M. Ziman, J. Phys. A **49**, 123001 (2016).
[27] R. Uola, C. Budroni, O. Gühne, and J. P. Pellonpää, Phys. Rev. Lett. **115**, 230402 (2015).
[28] P. Busch, Phys. Rev. D **33**, 2253 (1986).
[29] M. M. Wolf, D. Perez-Garcia, and C. Fernandez, Phys. Rev. Lett. **103**, 230402 (2009).

[30] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens, Europhys. Lett. **103**, 10002 (2013).
[31] M. Banik, M. R. Gazi, S. Ghosh, and G. Kar, Phys. Rev. A **87**, 052125 (2013).
[32] N. Stevens and P. Busch, Phys. Rev. A **89**, 022123 (2014).
[33] D. Cavalcanti and P. Skrzypczyk, Phys. Rev. A **93**, 052112 (2016).
[34] M. Plávala, Phys. Rev. A **94**, 042108 (2016).
[35] A. Jenčová and M. Plávala, Phys. Rev. A **96**, 022113 (2017).
[36] Y. Kuramochi, Positivity **24**, 1479 (2020).
[37] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.128.160402 for further details.
[38] H. Ollivier and W. H. Zurek, Phys. Rev. Lett. **88**, 017901 (2001).
[39] G. Adesso, T. R. Bromley, and M. Cianciaruso, J. Phys. A **49**, 473001 (2016).
[40] U. M. Maurer, IEEE Trans. Inf. Theory **39**, 733 (1993).
[41] M. Kläy, C. Randall, and D. Foulis, Int. J. Theor. Phys. **26**, 199 (1987).
[42] A. Wilce, Int. J. Theor. Phys. **31**, 1915 (1992).
[43] G. Aubrun, L. Lami, C. Palazuelos, and M. Plávala, Geom. Funct. Anal. **31**, 181 (2021).
[44] J. Clauser, M. Horne, A. Shimony, and R. Holt, Phys. Rev. Lett. **23**, 880 (1969).
[45] G. Barker, Linear Multilinear Algebra **4**, 191 (1976).
[46] G. Barker, Linear Algebra Appl. **39**, 263 (1981).
[47] L. Hardy, arXiv:quant-ph/0101012.
[48] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Phys. Rev. A **84**, 012311 (2011).
[49] L. Masanes and M. Müller, New J. Phys. **13**, 063001 (2011).
[50] A. J. P. Garner, M. P. Müller, and O. C. O. Dahlsten, Proc. R. Soc. A **473**, 20170596 (2017).
[51] H. Barnum and J. Hilgert, arXiv:1904.03753.
[52] H. Barnum and J. Hilgert, J. Lie Theory **30**, 315 (2020), https://www.heldermann.de/JLT/JLT30/JLT302/jlt30019.htm.
[53] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Phys. Rev. Lett. **99**, 240501 (2007).