Witnessing nonclassicality beyond quantum theory

Chiara Marletto and Vlatko Vedral

Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom;
Centre for Quantum Technologies, National University of Singapore,
3 Science Drive 2, 117543 Singapore, Singapore;
Department of Physics, National University of Singapore, 2 Science Drive 3, 117542 Singapore, Singapore;
and ISI Foundation, Via Chisola 5, 10126 Torino TO, Turin

(Received 9 August 2020; accepted 23 September 2020; published 16 October 2020)

We propose a general argument to show that if a physical system can mediate locally the generation of entanglement between two quantum systems, then it itself must be nonclassical. Remarkably, we do not assume any classical or quantum formalism to describe the mediating physical system: our result follows from general information-theoretic principles, drawn from the recently proposed constructor theory of information. This argument provides a broader theoretical basis for recently proposed tests of non-classicality in gravity, based on witnessing gravitationally induced entanglement in quantum probes, making them applicable in the context where quantum theory may not apply.

DOI: 10.1103/PhysRevD.102.086012

A class of experiments for detecting nonclassicality in gravity has recently been proposed [1,2]. This has opened up an exciting possibility: quantum effects in gravity can be detected by probing indirectly the nonclassicality of the gravitational interaction, through measuring the gravitationally induced entanglement on *two* quantum probes. In this paper, we focus on the theoretical foundations for experiments in this class.

These experiments are based on the fact that if a system $\mathbf{M}$ (e.g., gravity) can entangle two quantum systems $\mathbf{Q_A}$ and $\mathbf{Q_B}$ (e.g., two masses) by local interactions, then $\mathbf{M}$ must be nonclassical. By *nonclassical*, we mean, informally, that the mediator $\mathbf{M}$ must have at least two variables that cannot be measured to arbitrarily high accuracy simultaneously (i.e., by the same measuring system). This notion generalizes what in quantum theory is called "complementarity," and it will be defined formally later, in a way that does not rely on quantum theory's formalism.

When $\mathbf{M}$ is assumed to obey quantum theory, once can prove a special case of the above statement from theorems about local operations and classical communication [3]: a decoherent channel cannot entangle two other quantum systems by local operations. On this ground, one can say that if a channel can mediate entanglement, it cannot be a completely decoherent one. Now, in order to apply these theorems to the case of gravity, one has already to assume that it obeys quantum theory; an experiment based on this assumption would therefore test whether gravity has some coherence, which permits massive superpositions to be allowed beyond certain scales. The arguments in [1] and related proposals [4,5] follow this line of argument and generalize it to cases where the mediator's quantum observables are not measurable directly. Specifically, in [1], by "classical theory," the authors meant the decohered version of a specific quantum theory of gravity (linear quantum gravity). However, the proposed experiments aim to probe cases (such as gravity) where the mediator $\mathbf{M}$ *may or may not* obey quantum theory. Therefore, to provide an adequately broad theoretical foundation for the proposed tests, one needs to adopt less restrictive assumptions, without assuming quantum theory in full.

A more general argument in this spirit was proposed in [2,6], not assuming all the properties of quantum dynamics for the mediator. That argument, though, was still expressed via density operators, which are rooted in quantum theory's formalism. Here we provide a far more general argument in support of the proposed tests of nonclassicality, based on information-theoretic principles and the principle of locality *only* (to be defined precisely later). We will also define generalizations of concepts such as nonclassical and *observable* to describe the mediator, that are compatible with quantum theory's and general relativity's, but do not assume either of those theories or their formalism. Hence, the argument we propose here is more general in two ways: (i) it does not assume a specific dynamical law for the mediator, but only that it obeys two generally applicable principles and (ii) it does not assume quantum theory's formalism.

To achieve this generality, we resort to the principles of the constructor theory of information [7], which provide a useful guide when neither quantum theory nor general relativity can be assumed. These principles allow one not to assume any specific dynamics for the mediator, therefore making our approach more general than the existing hybrid quantum-classical approaches, such as [8,9], where a generalized Hamiltonian dynamics is usually assumed. Our logic here is akin to that of Bell's theorem: just like Bell's theorem applies to a vast class of theories obeying general probabilistic assumptions, our theorem applies to a set of theories that obey general information-theoretic principles, also in the spirit of other principle-based arguments proposed to merge quantum theory and general relativity, such as [10–12]. Note how the latter arguments, including [2], are far less general than the one presented in this paper, because they rely on assuming direct control over the system that has to be proven quantum, and they also assume the existence of specific measurementlike interactions. The arguments presented here instead assume only two principles, i.e., locality and interoperability of information, as defined in the next section. These principles are akin to say thermodynamic principles or conservation laws, which are not derived from dynamical laws, but instead are purported to constrain known laws and those yet to be known.

I. AN EXAMPLE FROM QUANTUM THEORY

This section discusses an example where the mediator does obey quantum theory to illustrate the logic of the general argument. We will use a qubit-based model, to elucidate the role of the mediator's nonclassical degrees of freedom in the entanglement generation. In this example, the relevant degrees of freedom are just the X and Z components of a mediating qubit $\mathbf{Q_M}$; in the general argument, we will show that the mediator must have analogous properties to $\mathbf{Q_M}$, but we shall prove this without assuming that $\mathbf{M}$ obeys quantum theory.

Consider two qubits, $\mathbf{Q_A}$ and $\mathbf{Q_B}$, and a mediator qubit $\mathbf{Q_M}$. They all start their evolution in a product state; at a later time, $\mathbf{Q_A}$ and $\mathbf{Q_B}$ become entangled, via interacting each *locally* with $\mathbf{Q_M}$. A simple model is an entangling gate acting between $\mathbf{Q_A}$ and $\mathbf{Q_M}$; then, a SWAP gate between $\mathbf{Q_M}$ and $\mathbf{Q_B}$. Assuming that $|0\rangle$ is the $+1$-eigenstate of the Z component of each qubit, an example of this entangling process is

$$|0\rangle_A|0\rangle_M|0\rangle_B \underbrace{\rightarrow}_{\mathrm{Bell_{AM}}} |B_{+0}\rangle_{AM}|0\rangle_B \underbrace{\rightarrow}_{\mathrm{SWAP_{MB}}} |B_{+0}\rangle_{AB}|0\rangle_M, \quad (1)$$

where $|B_{+0}\rangle$ is one of the Bell's states, describing two maximally entangled qubits. In order for $\mathbf{Q_A}$ and $\mathbf{Q_B}$ to become entangled via $\mathbf{Q_M}$, the latter must itself be entangled with $\mathbf{Q_A}$, at least to the same degree as qubits $\mathbf{Q_A}$ and $\mathbf{Q_B}$ are at the end of the protocol. This requires the mediator $\mathbf{Q_M}$ to engage other variables in its dynamical evolution, such as the X and Y components, which do not commute with its Z component (that it is initially sharp). Our proposed argument will establish the existence of a generalized version of these incompatible variables as the signature of nonclassicality of the mediator $\mathbf{M}$, without assuming that the latter obeys quantum theory.

We can see more clearly how the incompatible variables are engaged in the entanglement generation by considering the Heisenberg picture.

Let $q_{x\alpha}$ denote an operator representing the X component of qubit α (likewise for the Y and Z components). These operators are defined on the 2^3-dimensional Hilbert space of the three qubits. We have $q_{z\alpha}q_{x\alpha} = \mathrm{i}q_{y\alpha}$, $q_{z\alpha}^2 = id$ and likewise for all the other components, while components of different qubits commute. If the gate $U(t_n)$ operates between time t_n and t_{n+1}, we shall denote by

$$O_\alpha(t_{n+1}) = U(t_n)^\dagger O_\alpha(t_n) U(t_n), \quad (2)$$

the operator representing the observable O of system α after its action. The initial conditions are fixed by choosing particular values for $q_{x\alpha}(t_0)$, $q_{y\alpha}(t_0)$, $q_{z\alpha}(t_0)$, for all α's, and by the Heisenberg state ρ_H. We assume that the initial conditions are expressed as $q_{z\mathrm{A}}(t_0) = Z \otimes id^{\otimes 3} \equiv q_{z\mathrm{A}}$, where Z is a Pauli matrix, and so on. We choose the Heisenberg state to be $\rho_H = |\mathbf{0}\rangle\langle\mathbf{0}|$, the $+1$ eigenstate of the operator $\frac{1}{2}(id + Z)^{\otimes 3}$.

The state of each qubit α at time t is completely specified by at least two components, e.g., $\{q_{x\alpha}(t), q_{z\alpha}(t)\}$. In this picture, we can describe the entangling operation mentioned above, as causing the following dynamical evolutions on each of the three qubits:

$$\begin{aligned}
\mathbf{Q_A}&: \{q_{x\mathrm{A}}, q_{z\mathrm{A}}\} \rightarrow \{q_{z\mathrm{A}}q_{x\mathrm{M}}, q_{x\mathrm{A}}\} \rightarrow \{q_{z\mathrm{A}}q_{x\mathrm{M}}, q_{x\mathrm{A}}\}\\
\mathbf{Q_M}&: \{q_{x\mathrm{M}}, q_{z\mathrm{M}}\} \rightarrow \{q_{x\mathrm{M}}, q_{z\mathrm{M}}q_{x\mathrm{A}}\} \rightarrow \{q_{x\mathrm{B}}, q_{z\mathrm{B}}\}\\
\mathbf{Q_B}&: \{q_{x\mathrm{B}}, q_{z\mathrm{B}}\} \rightarrow \{q_{x\mathrm{B}}, q_{z\mathrm{B}}\} \rightarrow \{q_{x\mathrm{M}}, q_{z\mathrm{M}}q_{x\mathrm{A}}\}.
\end{aligned}$$

Here, the first column represents the initial value of the qubit's descriptors; the second column represents their values at time t_1, when qubits A and $\mathbf{M}$ are entangled; the third represents the final value, where qubits A and B are entangled: this can be verified by considering the expected value of an entanglement witness, evaluated at that time.

One can see that $\mathbf{Q_M}$'s two incompatible observables (its X and Z component) are both engaged in mediating, by local interactions, the quantum correlations between $\mathbf{Q_A}$ and $\mathbf{Q_B}$. The logic outlined here is widespread in quantum information and underpins protocols such as teleportation and entanglement swapping; but it is very useful to bear it in mind in view of our aim: we will establish that a general mediator $\mathbf{M}$ has to be nonclassical, just like $\mathbf{Q_M}$, without assuming that $\mathbf{M}$ obeys quantum theory. This will entail showing that in order to entangle two qubits by interacting with each one individually, **M** must have degrees of freedom that, analogously to the X and Z components of $\mathbf{Q_M}$ in the above example, are incompatible with each other. All of these notions will now be formally defined in this more general scenario where quantum theory may not fully be obeyed by the mediator.

II. THE INTEROPERABILITY PRINCIPLE FOR INFORMATION

Here we introduce a constructor-theoretic principle, the interoperability principle for information [7], and we express the principle of locality, which are the foundation of the argument we intend to propose. To this end, we will summarize the concepts of constructor theory needed in order to express those principles.

A. States, attributes, and variables

When dropping the assumption that a specific dynamics holds for **M**, we can still maintain other notions, such as a generalized notion of *state*—which provides the full description of a physical system. We will assume that **M** obeys a theory endowed with a set of allowed *states* for physical systems and a partition of the whole universe into subsystems. We will be concerned with physical systems on which transformations can be performed, called substrates.

An *attribute* **n** of a substrate is the set of all states where the substrate has a given property. A *variable* is a set of disjoint attributes of a substrate. (Note that variables and observables differ: the attributes in a variable may not be distinguishable, as explained below.)

A variable V is *sharp* on a given system, with value v, if the system is in a state belonging to the attribute **v** in that variable.

For instance, a qubit is a substrate; the set of all $+1$-eigenstates of a given projector is an attribute; that projector is sharp with value 1 whenever the qubit is in any one of those states.

B. Possible/impossible tasks

A *task* specifies a general physical transformation of a substrate, in terms of ordered pairs of input/output attributes. For example, the NOT task on the attributes **0**, **1** is written as $\{\mathbf{0} \to \mathbf{1}, \mathbf{1} \to \mathbf{0}\}$.

A task is *impossible* if the laws of physics impose a limit to how accurately it can be performed. Unitary quantum theory, for instance, requires the task of cloning sets of nonorthogonal quantum states to be impossible [13]. Otherwise, the task is *possible*: there can be arbitrarily good approximations to a *constructor* for it, which is defined as a substrate that, whenever presented with the substrates in any of the input attributes of the task, delivers them in one of the corresponding output attributes, and, crucially, retains the property of being capable of performing the task.

C. Locality as a constraint on states

A cardinal principle of constructor theory is the principle of locality, which can be expressed as a strict constraint on the states of substrates, as follows:

Principle 1: Locality.—The state of a substrate is a description of it that satisfies two properties: (i) any attribute of a substrate, at any given time t, is a *fixed function* of the substrate's *state* and (ii) any state of a composite substrate $\mathbf{S_1} \oplus \mathbf{S_2}$ is an ordered pair of states (s_1, s_2) of $\mathbf{S_1}$ and $\mathbf{S_2}$, with the property that if a task is performed on $\mathbf{S_1}$ only, then the state of the substrate $\mathbf{S_2}$ is not changed thereby.

The principle of locality in this form is satisfied by quantum theory, but the states do not correspond to the density operators. This is manifest by considering quantum theory's Heisenberg picture [14]. In the Heisenberg picture, the state of a quantum system is the vector of the generators of its algebra of observables (which are dynamical variables). For instance, in the case of a single qubit—in the notation introduced earlier—its state is the vector of time-dependent components $\hat{q} \doteq (q_x(t), q_y(t), q_z(t))$; the fixed function is $\mathrm{Tr}(\bullet\rho_0)$, where the dot stands for any appropriate set of Hermitian operators in the span of $\hat{q}$, and ρ_0 is some (fixed) Heisenberg state.

Now, considering a two-qubit system, the state of each qubit α at time t is completely specified by at least two components, e.g., $\{q_{x\alpha}(t), q_{z\alpha}(t)\}$. The state of the joint system is likewise reconstructed given all of the observables in the set $\{q_{x\alpha}(t),\ q_{z\alpha}(t)\}$, because

$$U(t_n)q_{x\alpha}(t_n)q_{z\alpha}(t_n)U^{\dagger}(t_n) = q_{x\alpha}(t_{n+1})q_{z\alpha}(t_{n+1}) \quad (3)$$

by unitarity. This is why quantum theory satisfies the principle of locality as expressed above, considering the q-valued descriptors of the Heisenberg picture as states. These descriptors are local in that sense because they contain all the information about a system's nontrivial history.

Note also that the principle of locality in this form implies no-signaling: for if the state of $\mathbf{S_2}$ does not change when a transformation on $\mathbf{S_1}$ is implemented, the empirically accessible attributes of $\mathbf{S_2}$ cannot change either, since, by the principle of locality, they are fully specified by a fixed function of that state [14,15].

D. Information media

One can provide a general information-theoretic characterization of the mediator **M** in our argument by resorting to the concept of information medium [7]. An *information medium* is a substrate with a set of attributes X, called *information variable*, with the property that the following tasks are possible:

$$\bigcup_{x\in X}\{(\mathbf{x},\mathbf{x_0})\to(\mathbf{x},\mathbf{x})\}, \tag{4}$$

$$\bigcup_{x\in X}\{\mathbf{x}\to\Pi(\mathbf{x})\} \tag{5}$$

for all permutation Π on the set of labels of the attributes in X and some blank attribute $\mathbf{x_0}\in X$.

The former task corresponds to "copying," or cloning, the attributes of the first replica of the substrate onto the second, target, substrate; the latter, for a particular Π, corresponds to a logically reversible computation (which need not require it to be realized in a physically reversible way). So, an information medium is a substrate that can be used for classical information processing (but could, in general, be used for more than just that). For example, a qubit is an information medium with respect to any set of two orthogonal quantum states.

E. The interoperability of information

Any two information media (e.g., a photon and an electron) must satisfy the principle of interoperability [7], which expresses the intuitive property that classical information must be copiable from one information medium to any other, irrespective of their physical details. Specifically:

Principle 2: If $\mathbf{S_1}$ and $\mathbf{S_2}$ are information media, respectively, with information variable X_1 and X_2, their composite system $\mathbf{S_1}\oplus\mathbf{S_2}$ is an information medium with information variable $X_1\times X_2$, where $\times$ denotes the Cartesian product of sets.

This principle implies that the task of copying information variables [as in Eq. (4)] from one information medium to the other is possible. It also requires the possibility of performing computations on $\mathbf{S_2}$ without simultaneously affecting $\mathbf{S_1}$; otherwise, it would not be possible to perform independent permutations of variables of $\mathbf{S_1}$ or $\mathbf{S_2}$. This property is guaranteed by the principle of locality, as expressed earlier.

We can now express information-theoretic concepts such as measuring and distinguishing, without resorting to formal properties such as orthogonality, linearity, or unitarity. This is the other key feature of constructor theory that will allow our argument to be independent of particular dynamical models. The variable X is *distinguishable* if the task

$$\bigcup_{x\in X}\{\mathbf{x}\to\mathbf{q_x}\} \tag{6}$$

is possible, where the variable $\{\mathbf{q_x}\}$ is some information variable. If the variable $\{\mathbf{x_0},\mathbf{x_1}\}$ is distinguishable, we say that the attribute $\mathbf{x_0}$ is distinguishable from $\mathbf{x_1}$. This notion of distinguishability allows one to generalize the orthogonal complement of a vector space: for any attribute $\mathbf{n}$, define the attribute $\bar{\mathbf{n}}$ as the union of all attributes that are distinguishable from $\mathbf{n}$.

An observable is an information variable whose attributes $\mathbf{x}$ have the property that $\bar{\bar{\mathbf{x}}}=\mathbf{x}$; this notion generalizes that of a quantum observable. An observable X is said to be sharp on a substrate, with value x, if the substrate is in a state that belongs to one of the attributes $\mathbf{x}\in X$.

A special case of the distinguishing task is the perfect measurement task,

$$\bigcup_{x\in X}\{(\mathbf{x},\mathbf{x_0})\to(\mathbf{x},\mathbf{p_x})\}, \tag{7}$$

where the first substrate is the "source" and the second substrate is the "target." From the interoperability principle, it follows that the above task must be possible for any information variable.

In the constructor theory of information, one can also define a generalization of quantum systems, called superinformation media [7]. A superinformation medium is an information medium with at least two information observables, X and Z, such that their union is not an information variable. We shall call these observables *incompatible*, borrowing the terminology from quantum theory, because a measurer of one must perturb a substrate where the other observable is sharp [7]. Qubits are special cases of superinformation media [7]: one can think of X and Z as two noncommuting observables, whose attributes cannot all be copied by the same cloner, because of the no-cloning theorem [13].

F. Nonclassicality

In our argument, we will aim at establishing that $\mathbf{M}$ has a lesser property: nonclassicality. By a substrate being nonclassical, we shall mean an information medium $\mathbf{M}$, with maximal information observable T, that has a variable V, disjoint from T and with the same cardinality as T, with the following properties:

1. There exists a superinformation medium $\mathbf{S_1}$ and a distinguishable variable $E=\{\mathbf{e_j}\}$ of the joint substrate $\mathbf{S_1}\oplus\mathbf{M}$, whose attributes $\mathbf{e_j}=\{(s_j,v_j)\}$ are sets of ordered pairs of states, where v_j is a state belonging to some attribute in V and s_j is a state of $\mathbf{S_1}$.
2. The union of V with T is *not* a distinguishable variable.
3. The task of distinguishing the variable $E=\{\mathbf{e_j}\}$ is possible by measuring incompatible observables of a *composite* superinformation medium including $\mathbf{S_1}$, but impossible by measuring observables of $\mathbf{S_1}$ only.

This generalizes the property of quantum complementarity to the case where $\mathbf{M}$ may not have the full information-processing power as a quantum system. For, contrary to superinformation media, in nonclassical substrates, the variable V may or may not be an information observable—it may not be permutable or copiable; yet, its existence requires $\mathbf{M}$ to enable nonclassical tasks on other superinformation media, such as establishing entanglement.

III. THE ARGUMENT

We can now formulate our central argument using these information-theoretic tools and principles, under the following assumptions:

(i) The mediator $\mathbf{M}$ is an information medium with a maximal information observable T.

(ii) The two systems to be entangled, $\mathbf{Q_A}$ and $\mathbf{Q_B}$, are qubits.

$\mathbf{Q_A}$ and $\mathbf{Q_B}$ qualify as superinformation media, having at least two *disjoint* maximal information observables, say their X and Z components, whose union is not an information observable. For simplicity, we will assume that all the information observables are binary: $T = \{\mathbf{t_0}, \mathbf{t_1}\}$; for the qubits, we have: $Z = \{\mathbf{z_1}, \mathbf{z_2}\}$ and $X = \{\mathbf{x_+}, \mathbf{x_-}\}$, where X and Z represent the X and Z components of each qubit, respectively.

We now proceed to demonstrate our main result.

Theorem 1. If $\mathbf{M}$ can entangle $\mathbf{Q_A}$ and $\mathbf{Q_B}$, by locally interacting with each, then $\mathbf{M}$ is nonclassical.

To prove this result, we will follow this logic. First, the interoperability principle implies that the following task is possible: to copy any of the observables of Q_α onto the observable T of the mediator $\mathbf{M}$, via some interaction. We will assume that by coupling M locally with each of the qubits via that same interaction, it is possible to prepare them in one of the two orthogonal maximally entangled states. By locality, this must be implemented by repeating two elementary steps: first performing a task on $\mathbf{Q_A} \oplus \mathbf{M}$ and then on $\mathbf{M} \oplus \mathbf{Q_B}$. We will run the argument assuming entanglement is obtained via these two elementary steps, as it is straightforward to generalize to the case where a repetition of the two steps is required. Upon performing the former task, $\mathbf{M}$ is prepared in one of the two attributes, by the principle of locality. These attributes, we shall argue, must belong to a binary variable V satisfying the nonclassicality conditions, just like the descriptors of the qubit Q_M in our qubit-based example.

We proceed now with presenting the argument in full. We first establish the fact that $\mathbf{Q_A} \oplus \mathbf{M}$ must have an additional variable E (generalizing a set of entangled states), as in the first condition for nonclassicality.

(i) Given the principle of interoperability, the task of measuring the observable X of one of the qubits, using the mediator $\mathbf{M}$ as the target, must be possible,

$$T_M \doteq \{(\mathbf{z_0}, \mathbf{t_0}) \to (\mathbf{z_0}, \mathbf{t_0}), \\ (\mathbf{z_1}, \mathbf{t_0}) \to (\mathbf{z_1}, \mathbf{t_1})\}, \qquad (8)$$

where the first slot represents one of the qubits; the second slot represents the mediator. In the limit of weak field, relevant for the tests in [1,2], one can think of $\mathbf{z_0}$ and $\mathbf{z_1}$ as two distinct locations of a mass, and of $\mathbf{t_0}$ and $\mathbf{t_1}$ as two distinguishable configurations of the gravitational field, induced by two different mass distributions $\mathbf{z_0}$ and $\mathbf{z_1}$. It is also possible to think of $\mathbf{t_0}$ and $\mathbf{t_1}$ as two distinguishable spacetime geometries, solutions of Einstein's equations for the two different mass distributions, as prescribed by general relativity [16].

(ii) If the experiment is successful in entangling $\mathbf{Q_A}$ and $\mathbf{Q_B}$, the following task must also possible:

$$T_E \doteq \{(\mathbf{x_+}, \mathbf{t_0}, \mathbf{x_+}) \to \mathbf{e_{++}}, \\ (\mathbf{x_-}, \mathbf{t_0}, \mathbf{x_+}) \to \mathbf{e_{-+}}\}, \qquad (9)$$

where $B \doteq \{\mathbf{e_{++}}, \mathbf{e_{-+}}\}$ is an information variable of $\mathbf{Q_A} \oplus \mathbf{M} \oplus \mathbf{Q_B}$ whose attributes correspond to two orthogonal, maximally entangled states of the two qubits. These attributes can be distinguished by measuring the observables of $\mathbf{Q_A}$ and $\mathbf{Q_B}$ only: specifically, let us assume that $\mathbf{e_{++}}$ is a maximally entangled state where both X_A, X_B and Z_A, Z_B are maximally correlated; while in $\mathbf{e_{-+}}$, the observables X_B and X_B are maximally correlated, while Z_A and Z_B are maximally anticorrelated. The proposed experiments [1,2] would show that the task T_E is possible, upon successfully generating entanglement between the probes $\mathbf{Q_A}$ and $\mathbf{Q_B}$.

(iii) Assume also that the constructor for the task T_E is the same as the constructor for the task T_M, so these two tasks can be performed jointly by the same interaction. In the case of the experiment with gravity, the constructor is the gravitational interaction between a mass and the gravitational field, initially prepared in some classical configuration, t_0. Also, we assume that T_E is performed without $\mathbf{Q_A}$ and $\mathbf{Q_B}$ interacting directly. By the principle of locality, it must be performed in at least two steps; the first only involving $\mathbf{Q_A}$ and $\mathbf{M}$, the second only $\mathbf{M}$ and $\mathbf{Q_B}$.

In the first step, this task is performed on $\mathbf{Q_A} \oplus \mathbf{M}$,

$$T_1 \doteq \{(\mathbf{x_+}, \mathbf{t_0}, \mathbf{x_+}) \to (\mathbf{s_{+0}}, \mathbf{x_+}), \\ (\mathbf{x_-}, \mathbf{t_0}, \mathbf{x_+}) \to (\mathbf{s_{-0}}, \mathbf{x_+})\}. \qquad (10)$$

In the second step, this other task on $\mathbf{M} \oplus \mathbf{Q_B}$ is performed,

$$T_2 \doteq \{(\mathbf{s_{+0}}, \mathbf{x_+}) \to \mathbf{e_{++}}, \\ (\mathbf{s_{-0}}, \mathbf{x_+}) \to \mathbf{e_{-+}}\}. \qquad (11)$$

From the possibility of task T_1, we see that the substrate $\mathbf{Q_A} \oplus \mathbf{M}$ has a variable: $E = \{\mathbf{s_{+0}}, \mathbf{s_{-0}}\}$. We now proceed to establish its properties to show that $\mathbf{M}$ is nonclassical.

(iv) First, note that E is a distinguishable variable, because it can be mapped one-to-one onto two distinguishable attributes of the qubits, $\mathbf{e}_{\alpha\beta}$, via task the T_2.

(v) By the principle of locality, there are states $\hat{q}_A^{\alpha 0}$ of $\mathbf{Q_A}$ and $m^{\alpha 0}$ of $\mathbf{M}$ such that each of the attributes in the variable $E = \{\mathbf{s}_{\alpha \mathbf{0}}\}$ is a fixed function of $(\hat{q}_A^{\alpha 0}, m^{\alpha 0})$ (where α takes values in $\{+,-\}$). Here $\hat{q}_A^{\alpha 0}$ is a vector of q-numbers representing the three components of the qubit, while $m^{\alpha 0}$ is some state describing $\mathbf{M}$, whose properties we wish to establish.

We proceed now to establish the properties of the set of attributes $V \doteq \{\{m^{\alpha 0}\}\}$ to show that $\mathbf{M}$ is nonclassical.

(1) *Condition 1 for nonclassicality.*—First, we prove that the set $V = \{\{m^{+0}\}, \{m^{-0}\}\}$ is a binary variable (i.e., a set of two disjoint attributes).

Proof.—The principle of locality requires the states $\mathbf{e}_{++}$ to be a fixed function of the states describing $\mathbf{Q_B}$ and $\mathbf{M}$ prior to performing T_2, likewise for $\mathbf{e}_{-+}$. Specifically, let us denote by $\hat{q}_B^{++}$ the state of $\mathbf{Q_B}$ *after* performing T_2, when the overall attribute is $\mathbf{e}_{++}$, and by $\hat{q}_B^{-+}$ the state of $\mathbf{Q_B}$ *after* performing T_2, when the overall attribute is $\mathbf{e}_{-+}$. By the principle of locality, $\hat{q}_B^{\alpha +} = H(\hat{q}_B, m^{\alpha 0})$, where H is some (well-behaved) function and $\hat{q}_B$ is a (q-numbered) state describing $\mathbf{Q_B}$ when it is in its initial attribute $\mathbf{x}_+$ (where X is sharp with value x_+).

We now use this fact to argue that $m^{+0} \neq m^{-0}$. First, e_{++} is distinguishable from e_{-+} only by measuring observables of *both* $\mathbf{Q_A}$ and $\mathbf{Q_B}$. Also, prior to performing T_2, the attributes $(\mathbf{s}_{+\mathbf{0}}, \mathbf{x}_+)$ and $(\mathbf{s}_{-\mathbf{0}}, \mathbf{x}_+)$, though overall distinguishable, are not distinguishable by measuring observables of $\mathbf{Q_B}$ jointly with observables of $\mathbf{Q_A}$. This is because $\mathbf{Q_B}$ is still in the same initial state $\hat{q}_B$ where the observable X is sharp with value x_+.

Thus, the state m^{+0} must be different from m^{-0}, as the dependence on $m^{\alpha 0}$ makes each of the $\{\hat{q}_B^{\alpha +}\}$ different from $\hat{q}_B$. Hence, the set of attributes $V = \{\{m^{+0}\}, \{m^{-0}\}\}$ is a variable (a set of disjoint attributes), with the same cardinality as T. Thus, $\mathbf{M}$ satisfies condition 1 for nonclassicality.

(2) *Condition 2 for nonclassicality.*—Next, we prove that the attributes in V are not distinguishable from, and do not overlap with, those in T.

Proof.—Given that the task $T_2 \cup T_M$ is possible (i.e., the two tasks are performed by the same constructor), each attribute $\{m^{\alpha 0}\}$ is not distinguishable from either $\mathbf{t_0}$ or $\mathbf{t_1}$. If it were, the attributes $\mathbf{x}_+$ and $\mathbf{x}_-$ of the qubit $\mathbf{Q_A}$ would be distinguishable from some of the $\mathbf{z}$'s, contrary to the assumption that $\mathbf{Q_A}$ is a superinformation medium. For the same reason, $m^{\alpha 0} \notin \mathbf{t_0}$ and $m^{\alpha 0} \notin \mathbf{t_1}$. Therefore, $\mathbf{M}$ satisfies condition (2) for nonclassicality.

(3) *Condition 3 for nonclassicality.*—We note that the variable V cannot be distinguished by measuring observables on $\mathbf{Q_A}$ only; it can be distinguished only by jointly measuring the complementary observables X_A and Z_A and X_B and Z_B on the superinformation medium $\mathbf{Q_A} \oplus \mathbf{Q_B}$. Hence, $\mathbf{M}$ satisfies also condition (3) for nonclassicality.

This concludes our proof that $\mathbf{M}$ is nonclassical.

IV. DISCUSSION

What could the attributes $\{m^{\alpha +}\}$ in the variable V be? Could they, for example, correspond to two different statistical mixtures of $\mathbf{M}$'s classical observable T, $\mathbf{t_0}$ and $\mathbf{t_1}$? The answer is no. This is because by performing the task T_2 and subsequently measuring observables of $\mathbf{Q_A}$ and $\mathbf{Q_B}$ jointly, one reveals entanglement between $\mathbf{Q_A}$ and $\mathbf{Q_B}$, which did not exist before the interaction between $\mathbf{Q_B}$ and $\mathbf{M}$. The correlations between observables of $\mathbf{Q_B}$ and those of $\mathbf{Q_A}$ after performing T_2 must be contained in the state $m^{\alpha +}$'s, given the locality principle: they are absent in $\mathbf{Q_B}$ before the interaction with $\mathbf{M}$, via T_2, while they are present in $\mathbf{Q_B}$ after performing T_2, when its state becomes dependent on $m^{\alpha +}$. Informally, the variable $V = \{\{m^{\alpha +}\}\}$ has at least the same information-carrying capacity as the q-number-valued states of the qubit $\mathbf{Q_B}$, because it contains all the correlations that are proper of an entangled qubit, as later confirmed by measurements of $\mathbf{Q_B}$. By Bell's theorem, $m^{\alpha +}$ cannot be a statistical mixture of $\mathbf{t_0}$ and $\mathbf{t_1}$, because, if it were, it would provide a local hidden variable model for quantum entangled states of $\mathbf{Q_A} \oplus \mathbf{Q_B}$. This argument, therefore, shows that collapse models, which would predict $\mathbf{M}$ to be in a statistical mixture of the observable T, are incompatible with observing entanglement.

Thus, the $\{m^{\alpha +}\}$ are not hidden variables, or "beables." They generalize the q-valued descriptors of what can dynamically change in a quantum system—the descriptors of the quantum Heisenberg picture. In this sense, they are closer to the observables as conceived by von Neumann in his argument to rule out hidden variable models [17]. Indeed, our argument could be understood as a first step toward generalizing Bell's theorem to inferring nonclassicality of systems, like $\mathbf{M}$, that can be used to *assist* locally the violation of Bell's inequalities on two other quantum systems, but need not have a full set of observables like a quantum system and therefore cannot violate Bell's inequalities directly.

Another interesting point is that the variable V may or may not be an information variable. If $\mathbf{M}$ were a qubit entangled with $\mathbf{Q_A}$, V could not be an information variable (otherwise we would be able to locally distinguish one entangled state from another just by measuring that information variable on $\mathbf{M}$). But given that system $\mathbf{M}$ may not obey quantum theory, so we must leave this possibility open. Note also that $\mathbf{M}$, although capable of working as a faithful channel for creating entanglement between the two qubits, may not have the full repertoire of operations such as preparation and measurement as a superinformation medium, let alone as a qubit.

Our argument does not commit to any particular formalism to describe **M** and its interaction with the two qubits, in contrast with the thorough analysis of the gravity experiment presented in [1,2,6,18], where specific dynamical models are assumed. But how general are the principles we assumed? The interoperability principle holds in any physical theory that allows for measurements and observables—whose existence is a prerequisite for any physical theory to be testable. Therefore, it is a robust principle. The principle of locality in the form discussed here is also satisfied by both quantum theory and general relativity. In [15], it also is proven that all theories based on 1∶1, no-signaling dynamics satisfy this principle of locality, thus making it a remarkably general property. This more general argument is of the essence for the witness of nonclassicality to hold irrespective of whether the mediator is assumed or not to obey specific quantum models. It is the essential theoretical underpinning for experiments assessing the quantization of gravity in full generality, where one cannot assume that gravity obeys a specific quantum model prior to the experiment. It ensures that if entanglement is observed, then *all* classical models for gravity, obeying our general principles, are ruled out. This is similar to Bell's theorem, which ensures that if Bell's inequalities are violated by a given theory, then all local hidden variable models for that theory are ruled out.

As also mentioned in [6], the observation of entanglement mediated by gravity, should it be achieved by an experiment, would only help us rule out classical theories for gravity. It would not give us, in the present form, information about confirming specific models of quantum gravity. Still, it would be interesting to analyze our proposed argument within existing quantum gravity theories, to understand what the nonclassical variable V is in each of the quantum gravity models that have been proposed, particularly nonperturbative ones. One could also consider lifting the assumption that $\mathbf{Q_A}$ and $\mathbf{Q_B}$ are qubits and proceed with the general theory of superinformation media [7,14,19], where entanglement is treated as locally inaccessible information. We conjecture that even in this case the degree of locally inaccessible information on $\mathbf{Q_A}$ and $\mathbf{Q_B}$ can be expressed formally as a function of the degree of nonclassicality of **M**, generalizing the formal relation existing in quantum theory between nonclassicality of the mediator and degree of entanglement [20].

This argument is effective to derive predictions in areas where specific dynamics cannot be assumed, going beyond current approximation schemes (such as open-system dynamics) or hybrid dynamical approaches (see e.g., [9]). The information-theoretic principles of constructor theory we used here provide a fruitful alternative to dynamics and initial conditions, useful to construct a bridge toward new theories of physics. In this paper, we have demonstrated the first experimental application of this powerful approach.

ACKNOWLEDGMENTS

The authors thank David Deutsch and Anicet Tibau Vidal for providing sharp criticism on this paper and an anonymous referee for providing helpful comments. C. M. thanks the Eutopia Foundation, the Templeton World Charity Foundation, the John Templeton Foundation, and the FQXi. V. V. thanks the National Research Foundation, Prime Minister's Office, Singapore, under its Competitive Research Programme (CRP Grant No. NRF-CRP14-2014-02) and administered by Centre for Quantum Technologies, National University of Singapore. This publication was made possible through the support of the ID 61466 grant from the John Templeton Foundation, as part of the Quantum Information Structure of Spacetime Project (qiss.fr). The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

[1] S. Bose, A. Mazumdar, G. W. Morley, H. Ulbricht, M. Toroš, M. Paternostro, A. A. Geraci, P. F. Barker, M. S. Kim, and G. Milburn, Phys. Rev. Lett. **119**, 240401 (2017).

[2] C. Marletto and V. Vedral, Phys. Rev. Lett. **119**, 240402 (2017).

[3] R. Horodecki, P. Horodecki, M. Horodecki, and K. Horodecki, Rev. Mod. Phys. **81**, 865 (2009).

[4] T. Krisnanda, M. Zuppardo, M. Paternostro, and T. Paterek, Phys. Rev. Lett. **119**, 120402 (2017).

[5] T. Krisnanda *et al.*, Quantum Inf. **6**, 12 (2020).

[6] C. Marletto and V. Vedral, Phys. Rev. D **98**, 046001 (2018).

[7] D. Deutsch and C. Marletto, Proc. R. Soc. A **471**, 2174 (2014).

[8] M. Hall, Phys. Rev. A **78**, 042104 (2008).

[9] A. Ahmadzadegan, R. Mann, and D. R. Terno, Phys. Rev. A **93**, 032122 (2016).

[10] C. Marletto and V. Vedral, npj Quantum Inf. **3**, 29 (2017).

[11] B. DeWitt, *A Global Approach to Quantum Field Theory*, International Series of Monographs on Physics (Oxford University Press, 2014).

[12] R. Feynman, *The Role of Gravitation in Physics: Report from the 1957 Chapel Hill Conference*, edited by C. M. DeWitt and D. Rickles (1957).

[13] W. K. Wootters and W. H. Zurek, Nature (London) **299**, 802 (1982).
[14] P. Heyden and D. Deutsch, Proc. R. Soc. A **456**, 1759 (2002).
[15] G. Brassard and P. Raymond-Robichaud, arXiv:1710.01380.
[16] M. Christodoulou and C. Rovelli, Phys. Rev. B **792**, 64 (2019).
[17] J. von Neumann, *Mathematical Foundations of Quantum Mechanics* (Princeton University Press, Princeton, 1955).
[18] R. Marshman *et al.*, arXiv:1907.01568.
[19] C. Marletto, Proc. R. Soc. A **472**, 2192 (2015).
[20] G. Bhole, J. A. Jones, C. Marletto, and V. Vedral, J. Phys. Commun. **4**, 025013 (2020).