No-Hypersignaling Principle

Michele Dall'Arno,[1,*] Sarah Brandsen,[1,†] Alessandro Tosini,[2,‡] Francesco Buscemi,[3,§] and Vlatko Vedral[4,1,∥]

[1]*Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, 117543 Singapore*
[2]*QUIT group, Physics Department, Pavia University, and INFN Sezione di Pavia, via Bassi 6, 27100 Pavia, Italy*
[3]*Graduate School of Informatics, Nagoya University, Chikusa-ku, 464-8601 Nagoya, Japan*
[4]*Atomic and Laser Physics, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX13PU, United Kingdom*

(Received 8 March 2017; published 14 July 2017)

A paramount topic in quantum foundations, rooted in the study of the Einstein-Podolsky-Rosen (EPR) paradox and Bell inequalities, is that of characterizing quantum theory in terms of the spacelike correlations it allows. Here, we show that to focus only on spacelike correlations is not enough: we explicitly construct a toy model theory that, while not contradicting classical and quantum theories at the level of spacelike correlations, still displays an anomalous behavior in its timelike correlations. We call this anomaly, quantified in terms of a specific communication game, the "hypersignaling" phenomena. We hence conclude that the "principle of quantumness," if it exists, cannot be found in spacelike correlations alone: nontrivial constraints need to be imposed also on timelike correlations, in order to exclude hypersignaling theories.

DOI: 10.1103/PhysRevLett.119.020401

One of the main tenets in modern physics is that if two spacelike separated events are correlated, then such correlations must not carry any information [1]. This assumption, constituting the so-called *no-signaling principle*, was the starting point used by Bell [2] to quantify and compare spacelike correlations of different theories on even grounds—an idea of vital importance for his argument about the Einstein-Podolsky-Rosen (EPR) paradox [3] and the derivation of his famous inequality. Subsequently, due to seminal works by Tsirelson (Cirel'son) [4] and Popescu and Rohrlich [5], it became clear that the no-signaling principle alone is not enough to characterize "physical" spacelike correlations: nonsignaling spacelike correlations allowed by quantum theory form a *strict* subset within the set of all nonsignaling correlations [6].

A natural question is then to try to identify additional principles that, together with the no-signaling principle, may be able to rule out all superquantum nonsignaling correlations at once. Various ideas have been proposed, ranging from complexity theory, e.g., the collapse of the complexity tower [7] to information theory, e.g., the information causality principle [8]. However, none of these has been able to characterize the quantum-superquantum boundary in full. In particular, an outstanding open question is whether quantum theory can be characterized in terms of the spacelike correlations it allows [6].

In this Letter, we show that this cannot be done: any approach to characterize quantum theory based only on spacelike correlations is necessarily incomplete unless it also takes into account timelike correlations as well. Our approach, which is completely unrelated to the study of temporal correlations in the manner of Leggett and Garg [9–12], considers the elementary resource of noiseless communication and the input-output correlations that can be so established. By analogy with the no-signaling principle, we operationally introduce what we call the "no-hypersignaling principle," which roughly states that any input-output correlation that can be obtained by transmitting a composite system should also be obtainable by independently transmitting its constituents. As obvious as this may look (it is indeed so in classical and quantum theories), the fact that quantum theory obeys the no-hypersignaling principle (as we define it) is in fact a highly nontrivial consequence of a recent result by Frenkel and Weiner [13]. We also notice that the no-hypersignaling principle is not related with phenomena such as superadditivity of capacities of noisy quantum channels [14].

We then construct a toy model theory, which violates the no-hypersignaling principle, but only possesses classical spacelike correlations. As such, this theory (and other analogous theories) would go undetected in any test involving only spacelike correlations, despite displaying the anomalous effect of hypersignaling. On the technical side, our model is closely related to the standard implementation [15–17] of Popescu–Rohrlich [5] superquantum nonsignaling spacelike correlations (or "PR boxes," for short). However, while the PR-box model theory relies on entangled states to outperform quantum spacelike correlations, our hypersignaling model relies on *entangled measurements* to outperform quantum timelike correlations. Nonetheless, since in our model only separable states are available, no superquantum spacelike correlation can be obtained. Therefore, while the standard PR-box model theory can be ruled out on the basis of its superquantum spacelike correlations, the model proposed here can only be ruled out by the principle of no-hypersignaling.

The no-hypersignaling principle.—In general, the starting point of a physical theory is to define its elementary systems. In *generalized probabilistic theories* (see Supplemental Material [18], and Refs. [19,20]), a system $S = (\mathcal{S}, \mathcal{E})$ is typically defined by giving a set of states $\mathcal{S}$ and a set of effects $\mathcal{E}$, representing, respectively, the preparations and the observations of the system. States can be arranged to form ensembles $\{\Omega_0, \Omega_1, \cdots\}$, and effects can be arranged to form measurements $\{E_0, E_1, \cdots\}$. The theory must also comprise a rule for computing the conditional probability of any effect on any state. For example, in quantum theory, a system is associated with a d-dimensional Hilbert space $\mathcal{H}$, states, and effects are represented by positive semidefinite operators on $\mathcal{H}$, and conditional probabilities are given by the Born (trace) rule. The theory must also include a set of transformations mapping states into states (or effects into effects): in the case of quantum theory, this is the set of quantum channels (i.e., completely positive and trace preserving linear maps).

Given an elementary system, an important role is played by its *dimension* [21], which is expected to depend solely on the set of states $\mathcal{S}$ and effects $\mathcal{E}$. Since one usually assumes that convex mixtures of states and effects can always be considered (following the idea that the randomization of different experimental setups is in itself another valid experiment), by linear extension, it is natural to introduce the real vector spaces $\mathcal{S}_{\mathbb{R}}$ and $\mathcal{E}_{\mathbb{R}}$, generated by real linear combinations of elements of $\mathcal{S}$ and $\mathcal{E}$, respectively. Notice that in typical situations, $\mathcal{E}_{\mathbb{R}}$ coincides with the set of linear functionals on $\mathcal{S}_{\mathbb{R}}$. One soon arrives at the following definition:

Definition 1. (linear dimension) The linear dimension of a system S, denoted by $\ell(S)$, is defined as the dimension of the real vector space $\mathcal{S}_{\mathbb{R}}$ (or $\mathcal{E}_{\mathbb{R}}$, which is the same in the finite dimensional case considered in this work).

The linear dimension of a classical system with d extremal states is equal to d, whereas the linear dimension of a quantum system associated with a d-dimensional Hilbert space is d^2. For convenience, we denote a d-dimensional classical system by C_d and a quantum system with d-dimensional Hilbert space by Q_d so that, in formula, $\ell(C_d) = d$ and $\ell(Q_d) = d^2$.

There are various ways proposed to make sense of this discrepancy: a typical solution is to define an "operational" dimension as the maximum number of states that can be distinguished in a single measurement, see, e.g., Ref. [22]. In this way, even though the linear dimension of a quantum system is d^2, the operational dimension turns out to be d, thus matching the dimension of the underlying Hilbert space. In what follows, we introduce an alternative operational definition of dimension, which is both widely applicable and is independent of any arbitrarily chosen task, such as perfect state discrimination.

In order to make our analysis more concrete, we need to introduce some notation. Given two finite alphabets $\mathcal{X} = \{x\}$ and $\mathcal{Y} = \{y\}$ containing m and n letters, respectively, let us consider the set of all m-input and n-output conditional probability distributions $p_{y|x}$ that can be generated by transmitting one elementary system S, when free shared randomness between sender and receiver is allowed. With this, we mean that the input x can be "encoded" on some ensemble $\{\Omega_x^{(\lambda)} : x \in \mathcal{X}\}$ while the output letter y is "decoded" whenever the corresponding outcome is obtained in some measurement $\{E_y^{(\lambda)} : y \in \mathcal{Y}\}$, where λ parametrizes the shared random variable. We denote the convex set of all such correlations by $\mathcal{P}_S^{m\to n}$. For example, $\mathcal{P}_{C_d}^{m\to n}$ is the set of all m-input and n-output conditional probability distributions that can be obtained by means of a d-dimensional classical noiseless channel and shared random data. Equivalently, $\mathcal{P}_{C_d}^{m\to n}$ can be characterized as the polytope whose vertices are exactly all those $p_{y|x}$ with either null or unit entries and such that $p_y \coloneqq \sum_x p_{y|x} \neq 0$ for at most d different values of x.

Crucial in our analysis is a recent result by Frenkel and Weiner [13], stating that, in the presence of shared classical randomness, any input-output correlation obtainable with a d-dimensional quantum system is also obtainable with a d-dimensional classical system (and vice versa)—in formula,

$$\mathcal{P}_{C_d}^{m\to n} = \mathcal{P}_{Q_d}^{m\to n},$$

for all (finite) values of m and n. We are thus motivated to introduce the following definition:

Definition 2. (signaling dimension) The signaling dimension of a system S, denoted by $\kappa(S)$, is defined as the smallest integer d such that $\mathcal{P}_S^{m\to n} \subseteq \mathcal{P}_{C_d}^{m\to n}$, for all m and n.

Note that $\kappa(S)$ equals the usual dimension, both in classical and quantum theories, and is thus a natural candidate for an operational definition of dimension. Moreover, $\kappa(S)$ only depends on the structure of $\mathcal{S}$ and $\mathcal{E}$, without relying on the (arbitrarily made) choice of any specific protocol such as state discrimination. Also, due to the already mentioned result of [13], in what follows, we will simply use the symbol $\mathcal{P}_d^{m\to n}$ to denote $\mathcal{P}_{C_d}^{m\to n}$, since the fact that the underlying theory is classical or quantum is immaterial for the problem at hand.

The no-hypersignaling principle is introduced by looking at how the dimension behaves under composition of elementary systems. In order to do this, we need the theory to provide us with a rule for combining multiple elementary systems into a larger one. For example, in quantum theory, the composition rule is given by the tensor product of the underlying Hilbert spaces. For the sake of the present Letter, we do not need to understand the various possible mechanisms with which elementary systems can be composed: given a set of elementary systems $\{S_k\}$, we denote their composition by $\otimes_k S_k$. Notice that the tensor product should here be interpreted only as a symbol denoting composition and is not necessarily related with the actual tensor product of vector spaces (the interested reader may refer to Ref. [23]).

However, it is natural to assume that the composition rule must satisfy some sensible constraints. For example, a first condition that must be met by any self-consistent theory is that any circuit obtained as the composition of systems, states, effects, and their transformations should produce non-negative conditional probabilities. An additional condition is that spacelike correlations obey the no-signaling principle, so that any instantaneous exchange of information is forbidden, see Fig. 1. There are still other, more subtle conditions that can be considered.

For example, Ref. [22] considers the condition of *local tomography*. This requires the state of a composite system to be determined by the statistics of measurements done independently on its constituents. This principle is not as obvious as that of no signaling; however, it arguably remains a sensible requirement for a theory that does not want to be "too holistic" (namely, the state of any composite systems should always be locally accessible). The principle of local tomography is related with the notion of dimension: a theory is locally tomographic whenever the linear dimension of a composite system does not exceed the product of the linear dimensions of its constituents, in formula,

$$\ell(\otimes_k S_k) \leq \prod_k \ell(S_k). \tag{1}$$

In fact, without resorting to truly exotic, *ad hoc* theories, the linear dimension of the composite system cannot be strictly less than the product of the linear dimensions of its constituents, so the inequality in Eq. (1) can be safely replaced with the equal sign (see Ref. [22] for further details on the concept of linear dimension).

The no-hypersignaling principle is the analogue of Eq. (1) stated for the signaling dimension, rather than the linear dimension. We thus have the following definition:

Definition 3. (No-hypersignaling principle) *A theory is nonhypersignaling if and only if, for any set of systems $\{S_k\}$ with signaling dimensions $\kappa(S_k)$, the signaling dimension of the composite system $\otimes_k S_k$ satisfies*

$$\kappa(\otimes_k S_k) \leq \prod_k \kappa(S_k). \tag{2}$$

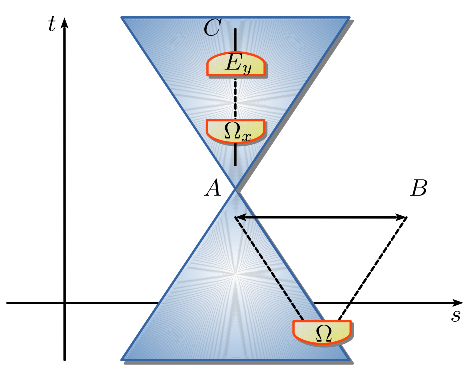

FIG. 1. *Spacelike and timelike correlations.* Events A and B are spacelike separated; i.e., information cannot travel from the one to the other (no-signaling principle). Correspondingly, they can only share spacelike correlations, previously distributed in the form of a bipartite state Ω. Events A and C are timelike separated; i.e., information can indeed travel from A to C: such information is encoded into the states $\{\Omega_x\}$, and later decoded by the measurement $\{E_y\}$. As the no-signaling principle constrains spacelike correlations, the no-hypersignaling principle constrains timelike correlations.

In particular, the no-hypersignaling principle requires that, given two copies of the same system S with signaling dimension d, the signaling dimension of $S \otimes S$ cannot exceed d^2, in formula

$$\mathcal{P}_S^{m\to n} \subseteq \mathcal{P}_d^{m\to n} \Rightarrow \mathcal{P}_{S^{\otimes 2}}^{m\to n} \subseteq \mathcal{P}_{d^2}^{m\to n},$$

for all m and n. The situation is depicted in Fig. 2.

Roughly speaking, while the no-signaling principle prevents spacelike separated parties from communicating, the no-hypersignaling principle prevents timelike separated parties from communicating "too much," see again Fig. 1. It may help to think that the no-hypersignaling principle guarantees that the input-output correlations, attainable when transmitting two elementary systems, do not depend on whether the systems are actually transmitted in series or in parallel.

Before proceeding, it will be useful to interpret the no-hypersignaling principle in terms of a communication game. To this aim, let us denote a composite system by $\bar{S} = \otimes_k S_k$ and by K the product $\prod_k \kappa(S_k)$ of the local signaling dimensions. It is therefore a straightforward application of the hyperplane separation theorem that a theory is hypersignaling if and only if, for some m and n, there exists a conditional probability distribution $p \in \mathcal{P}_{\bar{S}}^{m\to n}$ and an $m \times n$ real matrix g, such that

$$g^T \cdot p > \max_{q \in \mathcal{P}_K^{m\to n}} g^T \cdot q, \tag{3}$$

where we use the notation $A^T \cdot B$ to indicate the Hilbert-Schmidt dot product $\sum_{x,y} A_{x,y} B_{x,y} = \mathrm{Tr}[A^T B]$. Notice that the maximization problem in the rhs of Eq. (3) is in closed form: by linearity the maximum is attained on the vertices

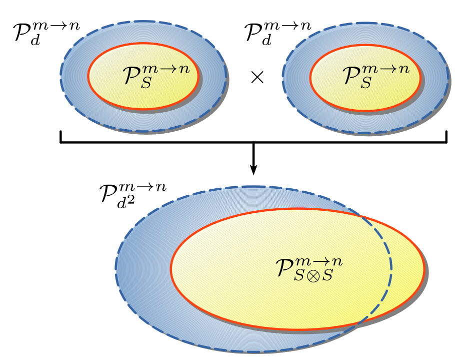

FIG. 2. *Illustration of a hypersignaling theory.* While the system S alone satisfies $\mathcal{P}_S^{m\to n} \subseteq \mathcal{P}_d^{m\to n}$, and thus has signaling dimension d, the composite system $S \otimes S$ has a signaling dimension strictly larger than d^2.

of the polytope $\mathcal{P}_K^{m\to n}$, which are finite in number and computed in the Supplemental Material [18].

The matrix g can be interpreted as the payoff function defining a communication game, where the sender inputs x and the receiver outputs y, leading to the corresponding payoff $g_{x,y}$. From this viewpoint, Eq. (3) represents the fact that, for any game g, the average payoff of the composite system $\bar{S}$ never exceeds the payoff of the product of its parts $\{S_k\}$. A general framework to consider such game-theoretic interpretation is developed in the Supplemental Material [18], by extending the theory of extremal quantum measurements [24,25] to general probabilistic theory.

The counterexample.—In what follows, we exploit our general framework to construct a toy model theory that violates the no-hypersignaling principle, namely such that the signaling dimension of the composite system is larger than the product of the signaling dimensions of its parts. Our toy model theory is explicitly derived along with all its constituents: elementary and composite systems, states, measurements, and dynamics. In the process, we clarify the relation between no-signaling, no-hypersignaling, local tomography, and information causality, arriving at the conclusion that the no-hypersignaling principle is independent of all of these and must therefore be assumed *separately.*

The elementary system here is the same as that used to reproduce PR correlations in Refs. [15–17]. The states and effects of the elementary system are vectors in $\mathbb{R}^3$, and there only exist four extremal states and four extremal effects, namely $\{\omega_x\}_{x=0}^3$ and $\{e_y\}_{y=0}^3$. As shown explicitly in the Supplemental Material [18] (see also Refs. [26,27]), all possible bipartite extensions can be given in terms of 24 extremal bipartite states, namely $\{\Omega_x\}_{x=0}^{23}$, and 24 extremal bipartite effects, namely $\{E_y\}_{y=0}^{23}$. The first 16 states (i.e., $0 \le x \le 15$) and the first 16 effects (i.e., $0 \le y \le 15$) are factorized, while the remaining ones are all entangled.

Because of self-consistency and the requirement that nontrivial reversible dynamics exist, however, bipartite states and effects cannot be chosen arbitrarily. As explicitly shown in the Supplemental Material [18], only the following three models satisfy all requirements: (i) *PR model* this is the theory used to model PR boxes [15–17]. It contains all possible extremal bipartite states, including the eight entangled ones (i.e., $\{\Omega_x\}_{x=0}^{23}$). Self-consistency then imposes that only extremal factorized effects are allowed (i.e., $\{E_y\}_{y=0}^{15}$). (ii) *HS model* this is the theory that we prove to be hypersignaling (HS). It contains only factorized extremal states (i.e., $\{\Omega_x\}_{x=0}^{15}$), but allows for all possible extremal effects, even entangled ones (i.e., $\{E_y\}_{y=0}^{23}$). (iii) *Hybrid models* in addition to all factorized states and effects, two entangled states and two entangled effects are allowed. Self-consistency singles out only two such models: states $\{\Omega_{20},\Omega_{22}\}$ with effects $\{E_{21},E_{23}\}$, or states $\{\Omega_{21},\Omega_{23}\}$ with effects $\{E_{20},E_{22}\}$.

Because of the presence of bipartite entangled states $\{\Omega_x\}_{x=16}^{23}$, the PR model is compatible with superquantum spacelike correlations, and this is actually the reason why it was introduced in the first place. However, we show in Supplemental Material [18], that the lack of entangled effects prevents the PR model from being hypersignaling. In a perfectly complementary way, the HS model cannot violate any Bell inequality, due to the lack of entangled states. However, in what follows, we show that, due to the presence of bipartite entangled effects $\{E_y\}_{x=16}^{23}$, the HS model violates the no-hypersignaling principle.

Let us start by noticing (see the Supplemental Material [18]) that the elementary system has a signaling dimension of two and is thus equivalent to the exchange of one classical bit. Therefore, to provide a counterexample to the no-hypersignaling principle, we need to provide a correlation ξ which is compatible with the composition of two elementary HS systems but cannot be obtained by exchanging only two classical bits.

One such a conditional probability has seven inputs and seven outputs and is given by

$$\xi = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 \end{pmatrix}. \tag{4}$$

This is explicitly obtained by applying the formalism developed in the Supplemental Material [18]. More explicitly, the rows of ξ are the conditional probabilities obtained by measuring the following measurement: $\{\frac{1}{8}E_0,\frac{1}{8}E_1,\frac{1}{8}E_6,\frac{1}{8}E_8,\frac{1}{8}E_{10},\frac{1}{8}E_{15},\frac{1}{4}E_{23}\}$, on each of the following seven states: $\{\Omega_0,\Omega_2,\Omega_6,\Omega_7,\Omega_{12},\Omega_{13},\Omega_{15}\}$.

The fact that ξ does not belong to $\mathcal{P}_4^{7\to 7}$, and thus violates the HS principle, is an immediate consequence of the characterization of polytope $\mathcal{P}_4^{7\to 7}$ provided in the Supplemental Material [18].

Since $\xi \notin \mathcal{P}_4^{7\to 7}$, there exists a game which violates Eq. (3). Indeed, consider the following game matrix g:

$$g = \frac{1}{21}\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 2 & 0 & 0 & 0 & 0 & 2 \\ 0 & 2 & 2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & 1 & 0 \end{pmatrix}.$$

It immediately follows by explicit computation that $g^T \cdot \xi = \frac{1}{2}$, while $\max_{q\in\mathcal{P}_4^{7\to7}} g^T \cdot q = \frac{10}{21} < \frac{1}{2}$. This latter result can be verified by explicitly computing the payoff associated with game g for all of the vertices of the polytope $\mathcal{P}_4^{7\to7}$, which are 35 9863 in number as shown in the Supplemental Material [18]. The interested reader can play the game of selecting four columns of g and further selecting one entry per row (within these columns), with the aim of maximizing the sum of the selected entries. They will then verify that no strategy will lead to a payoff larger than $\frac{10}{21}$.

Outlooks.—We have seen how it is possible to construct a generalized probabilistic theory—the HS model—that contradicts quantum theory, but only in timelike scenarios. This is consequence of the fact that the HS model has been arranged so that only separable states are allowed. In this way, when measurements are restricted to be separable due to locality constraints (as it is the case when testing spacelike correlations), the HS model never goes beyond classical theory. However, the possibility of having entangled measurements enables hypersignaling, thus proving that the HS model indeed goes *beyond* quantum theory in timelike scenarios.

It is now important to understand how hypersignaling is logically related with other possible "anomalies," such as the violation of local tomography or the violation of information causality. If any hypersignaling theory necessarily violates also other principles concerning spacelike correlations, then one could rightly argue that the phenomenon of hypersignaling might be ruled out just by looking at spacelike correlations. However, the point of this Letter is to argue the opposite: that timelike correlations require a new *independent* principle.

The fact that hypersignaling and information causality are independent is easy to see. As a necessary condition for the violation of information causality is the presence of entangled states, and since the HS model only contains separable states, then the HS model necessarily obeys information causality, despite allowing hypersignaling. Vice versa, we know that the PR model violates information causality but, since it only allows separable measurements, it cannot display any form of hypersignaling. The situation is depicted in left Fig. 3.

We now turn to the condition of local tomography [22]. From the explicit expression of the pure states of the HS model, it is possible to verify, as done in the Supplemental Material [18], that the elementary system S has linear dimension $\ell(S) = 3$ and that the bipartite system $S \otimes S$ has linear dimension $\ell(S \otimes S) = 9 = \ell(S)^2$. Thus, the HS model is locally tomographic, despite being hypersignaling. Vice versa, there exist consistent theories that obey the no-hypersignaling principle and yet are not locally tomographic. As an example, let us consider restrictions (for example, superselections) of quantum theory, as introduced in Ref. [28]. Since such theories are restrictions of quantum

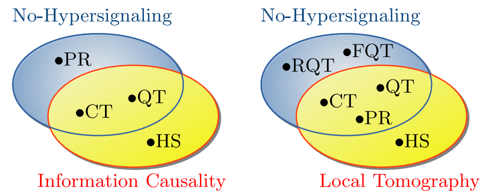

FIG. 3. *No-hypersignaling vs information causality and vs local tomography.* Left: the diagram compares theories satisfying information causality (yellow set) and the no-hypersignaling principle (blue set): CT (classical theory), QT (quantum theory), PR model (the toy model theory for PR boxes), and HS model (the locally classical, hypersignaling theory constructed in this Letter). Right: comparison between local tomography and no hypersignaling as two features of general probabilistic theories. Examples of theories that are nonhypersignaling but violate local tomography are provided by real quantum theory (RQT) and fermionic quantum theory (FQT). The HS model is locally tomographic but hypersignaling. Finally CT, QT, and the PR model lie in the intersection, as they obey both local tomography and the no-hypersignaling principle.

theory, they cannot exhibit hypersignaling: if they did, then quantum theory would also exhibit hypersignaling, which is not true. For example, real quantum theory [22] and fermionic quantum theory [28] are two possible such restricted quantum theories. However, as proved in Refs. [22,28,29], both theories are not locally tomographic. The situation is summarized in right Fig. 3.

We also notice that the no-hypersignaling principle can be violated by theories that do not show superadditivity of classical capacities. In Ref. [30], the authors show that a locally tomographic theory cannot feature superadditivity effects of classical capacities. Thus, hypersignaling does not necessarily imply superadditivity of classical capacities, because the HS model is locally tomographic. In passing by, the maximal mutual information for the hypersignaling correlation ξ in Eq. (4) (numerically optimized over any prior probability distribution) is less than 1.78 bits, which is below the Holevo bound of $\log_2 4 = 2$.

One interesting question arises from noting that while the HS model has classical spacelike correlations and superquantum timelike correlations, the PR model has superquantum spacelike correlations and classical timelike correlations. Could it be that a theory can be superquantum only with respect to either spacelike or timelike correlations, but not both? Could quantum theory have the unique distinction of "balancing" between these two extrema? It turns out that the answer is no, and follows from the example of the hybrid models derived above. In order to obtain the hypersignaling ξ in Eq. (4), we need seven factorized states and seven effects among which only one, precisely E_{23}, is not factorized. Since E_{23} is exactly one of those entangled effects admitted in the hybrid models, we know that the same ξ can be surely obtained in those models too. Moreover, since in the hybrid models two entangled states are also available, superquantum spacelike correlations can also be created. Hence, the hybrid models have the ability to create both spacelike and timelike superquantum correlations.

Finally, we compare the no-hypersignaling principle with two recently proposed and related principles, that is, *dimension mismatch* [21] and *information content* [31]. Both such principles rule out superquantum theories on the basis of the correlations achievable by a single-partite system, in contrast with the no-hypersignaling principle which requires composite systems. However, they achieve this by considering a more complicated setup, where the choice of the information to be decoded is not fixed but depends on an additional input (a second question) to the receiver. Moreover, both the dimension mismatch principle and the information content principle rely on a certain degree of arbitrariness in the criteria chosen to benchmark operational theories: dimension mismatch is defined with respect to an arbitrarily chosen reference task, i.e., pairwise state discrimination, while information content is defined with respect to an arbitrarily chosen information measure, i.e., mutual information. This is in contrast with the no-hypersignaling principle proposed here, where the full set of input-output correlations is considered without the need to invoke any particular discrimination task or information measure.

M. D. acknowledges support from the Singapore Ministry of Education Academic Research Fund Tier 3 (Grant No. MOE2012-T3-1-009). A. T. acknowledges support from the Templeton Foundation, Grant No. 60609. F. B. acknowledges support from the Japan Society for the Promotion of Science (JSPS) KAKENHI, Grant No. 17K17796. V. V. acknowledges support from the Ministry of Education and the Ministry of Manpower (Singapore).

*cqtmda@nus.edu.sg
†sbrandse@caltech.edu
‡alessandro.tosini@unipv.it
§buscemi@is.nagoya-u.ac.jp
‖phyvv@nus.edu.sg

[1] A. Einstein, Zur Elektrodynamik bewegter Körper, Ann. Phys. (Berlin) **17**, 891 (1905).
[2] J. Bell, On the Einstein Podolsky Rosen paradox, Physics (Long Island City, N.Y.) **1**, 195 (1964); Reprinted in J. S. Bell, *Speakable and Unspeakable in Quantum Mechanics* (Cambridge University Press, Cambridge, 1987), p. 14.
[3] A. Einstein, B. Podolsky, and N. Rosen, Can quantum-mechanical description of physical reality be considered complete?, Phys. Rev. **47**, 777 (1935).
[4] B. S. Tsirelson, On the Einstein Podolsky Rosen paradox, Lett. Math. Phys. **4**, 93 (1980).
[5] S. Popescu and D. Rohrlich, Quantum nonlocality as an axiom, Found. Phys. **24**, 379 (1994).
[6] S. Popescu, Nonlocality beyond quantum mechanics, Nat. Phys. **10**, 264 (2014).
[7] W. van Dam, Implausible consequences of superstrong nonlocality, Nat. Comput. **12**, 9 (2013).
[8] M. Pawłowski, T. Paterek, D. Kaszlikowski, V. Scarani, A. Winter, and M. Żukowski, Information causality as a physical principle, Nature (London) **461**, 1101 (2009).
[9] S. Brierley, A. Kosowski, M. Markiewicz, T. Paterek, and A. Przysiezna, Nonclassicality of Temporal Correlations, Phys. Rev. Lett. **115**, 120404 (2015).
[10] C. Budroni and C. Emary, Temporal Quantum Correlations and Leggett-Garg Inequalities in Multilevel Systems, Phys. Rev. Lett. **113**, 050401 (2014).
[11] C. Budroni, T. Moroder, M. Kleinmann, and O. Gühne, Bounding Temporal Quantum Correlations, Phys. Rev. Lett. **111**, 020403 (2013).
[12] M. Markiewicz, P. Kurzyński, J. Thompson, S.-Y. Lee, A. Soeda, T. Paterek, and D. Kaszlikowski, Unified approach to contextuality, nonlocality, and temporal correlations, Phys. Rev. A **89**, 042109 (2014).
[13] P. E. Frenkel and M. Weiner, Classical information storage in an n-level quantum system, Commun. Math. Phys. **340**, 563 (2015).
[14] M. B. Hastings, Superadditivity of communication capacity using entangled inputs, Nat. Phys. **5**, 255 (2009).
[15] J. Barrett, N. Linden, S. Massar, S. Pironio, S. Popescu, and D. Roberts, Nonlocal correlations as an information-theoretic resource, Phys. Rev. A **71**, 022101 (2005).
[16] J. Barrett, Information processing in generalized probabilistic theories, Phys. Rev. A **75**, 032304 (2007).
[17] G. M. D'Ariano and A. Tosini, Testing axioms for quantum theory on probabilistic toy-theories, Quantum Inf. Process. **9**, 95 (2010).
[18] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.119.020401 for details on the technical framework developed to derive the results presented in this Letter.
[19] R. W. Spekkens, Evidence for the epistemic view of quantum states: A toy theory, Phys. Rev. A **75**, 032110 (2007).
[20] G. Chiribella, G. M. D'Ariano, and P. Perinotti, Informational derivation of quantum theory, Phys. Rev. A **84**, 012311 (2011).
[21] N. Brunner, M. Kaplan, A. Leverrier, and P. Skrzypczyk, Dimension of physical systems, information processing, and thermodynamics, New J. Phys. **16**, 123050 (2014).
[22] L. Hardy and W. K. Wootters, Limited holism and real-vector-space quantum theory, Found. Phys. **42**, 454 (2012).
[23] H. Barnum, J. Barrett, M. Leifer, and A. Wilce, Teleportation in general probabilistic theories, in *Proceedings of Symposia in Applied Mathematics*, edited by S. Abramsky and M. Mislove (American Mathematical Society, Providence, RI, 2012), Vol. 71, p. 25.
[24] K. R. Parthasarathy, Extremal decision rules in quantum hypothesis testing, Inf. Dim. Anal. Quant. Prob. Rel. Topics **02**, 557 (1999).
[25] G. M. D'Ariano, P. Lo Presti, and P. Perinotti, Classical randomness in quantum measurements, J. Phys. A **38**, 5979 (2005).
[26] 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).

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

[28] G. M. D'Ariano, F. Manessi, P. Perinotti, and A. Tosini, The Feynman problem and Fermionic entanglement: Fermionic theory versus qubit theory, Int. J. Mod. Phys. A **29**, 1430025 (2014).

[29] G. M. D'Ariano, F. Manessi, P. Perinotti, and A. Tosini, Fermionic computation is non-local tomographic and violates monogamy of entanglement, Europhys. Lett. **107**, 20009 (2014).

[30] S. Massar, S. Pironio, and D. Pitalúa-García, Hyperdense coding and superadditivity of classical capacities in hypersphere theories, New J. Phys. **17**, 113002 (2015).

[31] L. Czekaj, M. Horodecki, P. Horodecki, and R. Horodecki, Information content of systems as a physical principle, Phys. Rev. A **95**, 022119 (2017).