

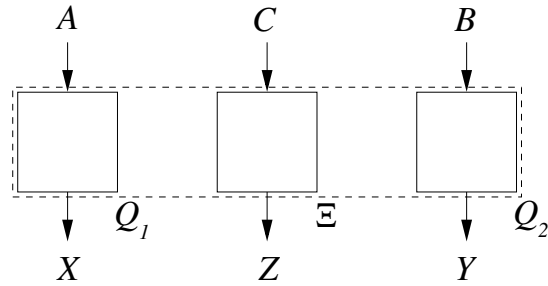
No extension of quantum theory can have improved predictive power

Supplementary Information

Roger Colbeck^{1,*} and Renato Renner²

¹*Perimeter Institute for Theoretical Physics, 31 Caroline Street North, Waterloo, ON N2L 2Y5, Canada.*

²*Institute for Theoretical Physics, ETH Zurich, 8093 Zurich, Switzerland.*



Supplementary Figure S1 | Abstraction of the setup. Q_1 and Q_2 depict a pair of quantum systems with inputs A and B and outputs X and Y respectively. E is a system which represents the additional information provided by the extended theory. Although these three systems (solid boxes) can be independently manipulated, they form parts of a larger system (dotted box). While no restriction is placed on the internal behaviour of the larger system, it follows from Part I of the proof that the combined distribution, $P_{XYZ|ABC}$, is non-signalling.

SUPPLEMENTARY METHODS

FORMAL STATEMENT OF THE CLAIM

In this section, we provide a formal description of our result and the assumptions it is based on. Their physical significance is explained in the main text.

Definitions

Definition 1. A *spacetime random variable* (SV), X , is a random variable¹ together with a set of coordinates $(t, r_1, r_2, r_3) \in \mathbb{R}^4$.

The coordinates can be used to define an order relation between SVs, which one may interpret as a time ordering within relativistic spacetime. (Note, however, that on a formal level, we do not require any assumptions about relativity theory.)

Definition 2. We say that a pair (A, X) of SVs is *time-ordered*, denoted $A \rightsquigarrow X$, if the coordinate (t, r_1, r_2, r_3) of A lies in the backward lightcone of the coordinate (t', r'_1, r'_2, r'_3) of X , i.e., $(t - t')^2 \geq \|r - r'\|^2$, $t \leq t'$. Furthermore, we say that two time-ordered pairs $A \rightsquigarrow X$ and $B \rightsquigarrow Y$ are *spacelike separated* if $A \not\rightsquigarrow Y$ and $B \not\rightsquigarrow X$.

The next two definitions refer to quantum theory or, more precisely, quantum measurements. They will be used later for the formulation of Assumption *QM*.

Definition 3. A *quantum measurement*, denoted $(A \rightsquigarrow X, \{E_x^a\}_{a,x}, \mathcal{H}_S)$, is a pair of time-ordered SVs, $A \rightsquigarrow X$, called *input* and *output*, respectively, together with a family of measurement operators $\{E_x^a\}_{a,x}$ on a Hilbert space, \mathcal{H}_S , such that $\sum_x (E_x^a)^\dagger E_x^a = \mathbb{1}_S$ for all a .

We interpret the input A as the choice of an observable and X as the outcome of the measurement with respect to this observable. Quantum theory determines the distribution of X conditioned on A , depending on the quantum state ρ_S of the system to which the measurement is applied.

Definition 4. Given a density operator ρ_S on \mathcal{H}_S , the quantum measurement $(A \rightsquigarrow X, \{E_x^a\}_{a,x}, \mathcal{H}_S)$ is said to be *compatible* with ρ_S if

$$P_{X|A}(x|a) = \text{tr}((E_x^a)^\dagger E_x^a \rho_S),$$

for all a and x . Likewise, a pair of quantum measurements $(A \rightsquigarrow X, \{E_x^a\}_{a,x}, \mathcal{H}_S)$ and $(B \rightsquigarrow Y, \{F_y^b\}_{b,y}, \mathcal{H}_T)$ is said to be *compatible* with ρ_{ST} defined on $\mathcal{H}_S \otimes \mathcal{H}_T$ if

$$P_{XY|AB}(xy|ab) = \text{tr}(((E_x^a)^\dagger E_x^a \otimes (F_y^b)^\dagger F_y^b) \rho_{ST}),$$

for all a, b, x and y .

We describe the process of *choosing* a value A as a pair of SVs, $O_A \rightsquigarrow A$, where O_A is called the *trigger event* (O_A may be a constant). The process is considered *free* if the outcome A is not correlated to anything that existed before the trigger event O_A in any reference frame.

Definition 5. Given a set of SVs Γ , a *free choice* (with respect to Γ) is a pair of time-ordered SVs, $O_A \rightsquigarrow A$, such that A is statistically independent of the collection $\Gamma' := \{W \in \Gamma : O_A \not\rightsquigarrow W\}$, i.e., $P_{A\Gamma'} = P_A \times P_{\Gamma'}$.

Quantum-Mechanical Description of the Measurement Process

Before stating our assumptions, let us briefly recall the quantum-mechanical description of a measurement process. Most generally, a quantum measurement on a system S is described by a family $\{E_x\}_x$ of operators acting on a Hilbert

¹ A note on notation: we will use upper case to represent random variables or SVs and lower case to represent specific values taken by such variables.

space \mathcal{H}_S such that $\sum_x E_x^\dagger E_x = \mathbb{1}$. If the state of S before the measurement is given by a density operator ρ_S then each possible outcome $X = x$ has probability

$$P_X(x) = \text{tr}(E_x^\dagger E_x \rho_S) .$$

(Note that this is reflected by Definitions 3 and 4.) Furthermore, conditioned on this outcome, the state of S after the measurement is

$$\sigma_S^{(x)} = \frac{E_x \rho_S E_x^\dagger}{P_X(x)} .$$

Averaged over all outcomes, the state is therefore given by $\sigma_S = \mathcal{E}(\rho_S)$, where \mathcal{E} is the trace-preserving completely positive map (TPCPM) defined by

$$\mathcal{E} : \rho_S \mapsto \sigma_S = \sum_x P_X(x) \sigma_S^{(x)} = \sum_x E_x \rho_S E_x^\dagger .$$

The TPCPM \mathcal{E} can be seen as part of an extended TPCPM $\bar{\mathcal{E}} : \rho_S \mapsto \sigma_{SDR}$ (in the sense that $\mathcal{E} = \text{tr}_{DR} \circ \bar{\mathcal{E}}$) which specifies the joint state σ_{SDR} of S , the measurement device, D , and possibly (parts of) the environment, R , after the measurement (one may think of $\bar{\mathcal{E}}$ as describing the joint evolution that the system S , measurement device D and the environment R undergo during a measurement). By choosing a sufficiently large environment, we can always take $\bar{\mathcal{E}}$ to be an isometry. Since the measurement outcome X is determined by the final state of the measurement device D , there exists a family of mutually orthogonal projectors $\{\Pi_x\}_x$ on the associated Hilbert space \mathcal{H}_D , where each Π_x projects onto the subspace containing the support of the state of D corresponding to outcome $X = x$. Formally, this corresponds to the requirement that

$$\forall x : \text{tr}_{DR}[\bar{\mathcal{E}}(\rho_S)(\mathbb{1}_S \otimes \Pi_x \otimes \mathbb{1}_R)] = E_x \rho_S E_x^\dagger . \quad (\text{S1})$$

Assumptions

To formulate our assumptions as well as our main claim, we consider an arbitrary quantum measurement

$$(A \rightsquigarrow X, \{E_x^a\}_{a,x}, \mathcal{H}_S) \quad (\text{S2})$$

with constant input $A = \bar{a}$ and output X . Furthermore, we consider two SVs, C and Z , such that $C \rightsquigarrow Z$, which model the access to extra information provided by a potential extended theory.

Our first assumption demands that the measurement we consider is correctly described by quantum mechanics.

Assumption QMa. *There exists a pure quantum state ρ_S which is compatible with the quantum measurement (S2).*

For the next assumption, let $\bar{\mathcal{E}} : \rho_S \mapsto \sigma_{SDR}$ be an isometry from \mathcal{H}_S to $\mathcal{H}_S \otimes \mathcal{H}_D \otimes \mathcal{H}_R$ and let $\{\Pi_x\}_x$ be a family of projectors such that (S1) holds for the operators $\{E_x^a\}_x$ specified by the measurement (S2).² The isometry $\bar{\mathcal{E}}$ models the joint evolution of the system, S , on which the measurement (S2) is carried out, the measurement device, D , and the parts of the environment, R , that may have been affected by the measurement.³ We then consider arbitrary measurements $\{F_x^a\}_{a,x}$ and $\{G_y^b\}_{b,y}$ on the subsystems D and SR , respectively, with the property that $F_x^{\bar{a}} = \Pi_x$.

The following assumption demands that the statistics produced by these additional measurements are as predicted by quantum theory. Furthermore, the outcome X of the initial measurement (S2) can be recovered by measuring (in an appropriate basis) the state of the device D used for this measurement.

Assumption QMb. *For appropriately defined SVs A', X', B, Y , the quantum measurements $(A' \rightsquigarrow X', \{F_x^a\}_{a,x}, \mathcal{H}_D)$ and $(B \rightsquigarrow Y, \{G_y^b\}_{b,y}, \mathcal{H}_S \otimes \mathcal{H}_R)$ are compatible with $\sigma_{SDR} = \bar{\mathcal{E}}(\rho_S)$. Furthermore, the measurement on D is consistent with the initial measurement (S2), in the sense that $X' = X$ whenever $A' = A = \bar{a}$.*

² There are many ways to choose $\bar{\mathcal{E}}$ and $\{\Pi_x\}_x$ with this property; our next assumption need only hold for one such choice.

³ Note that, for (S1) to hold, it is sufficient that $\bar{\mathcal{E}}$ describes the interaction between S and D (and possibly R) on a microscopic scale and for a short time. Hence, the fact that $\bar{\mathcal{E}}$ is an isometry does not preclude subsequent ‘‘collapse’’ of the wave function.

While the above assumptions are essentially consequences of the requirement that the existing quantum theory is correct, our last assumption demands that the measurement settings can be chosen freely.

Assumption FR. *There exist SVs O_A , O_B and O_C with $O_A \rightsquigarrow X'$, $O_B \rightsquigarrow Y$ and $O_C \rightsquigarrow Z$ spacelike separated such that $O_A \rightsquigarrow A'$, $O_B \rightsquigarrow B$ and $O_C \rightsquigarrow C$ are free choices with respect to $\{A', B, C, X', Y, Z\}$, and all possible values of A' and B are taken with nonzero probability.*

Main Claim

Theorem 1. If the quantum measurement (S2), modelled by the pair $A \rightsquigarrow X$, and the additional information, $C \rightsquigarrow Z$, are such that Assumptions QM_a , QM_b and FR are satisfied then the Markov chain condition $X \leftrightarrow (A, C) \leftrightarrow Z$ holds.

PART II OF THE PROOF

In this section, we prove the core inequality of Part II of our proof, Eqn. 8 in the Methods, which is stated as Lemma 1 below.

Recall the bipartite scenario described in the main text. The measurements at each site are parameterized by values $A \in \{0, 2, \dots, 2N - 2\}$ and $B \in \{1, 3, \dots, 2N - 1\}$ for some $N \in \mathbb{N}$, and their respective outcomes, X and Y , are taken to be binary. The measurements give rise to a joint probability distribution $P_{XY|AB}$ from which we quantify the correlations relevant for our statement in terms of I_N defined by

$$I_N(P_{XY|AB}) := P(X = Y|A = 0, B = 2N - 1) + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y|A = a, B = b) .$$

We consider enlargements of this probability distribution, $P_{XYZ|ABC}$ (see Figure S1), that satisfy the non-signalling property (cf. Part I of the proof), i.e.,

$$P_{XY|ABC} = P_{XY|AB} \tag{S3}$$

$$P_{XZ|ABC} = P_{XZ|AC} \tag{S4}$$

$$P_{YZ|ABC} = P_{YZ|BC} . \tag{S5}$$

The claim is that any such extension approximately satisfies $P_{Z|abcx} = P_{Z|abc}$, i.e., Z is independent of X for any choices of a , b and c . The accuracy of the approximation is measured in terms of the variational distance. For two distributions, P_X and P_Y over identical alphabets, this is defined by $D(P_X, P_Y) := \frac{1}{2} \sum_i |P_X(i) - P_Y(i)|$.

Lemma 1. *For any non-signalling probability distribution, $P_{XYZ|ABC}$, we have*

$$D(P_{Z|abcx}, P_{Z|abc}) \leq I_N(P_{XY|AB}) \tag{S6}$$

for all a , b , c , and x .

The proof is a generalization of an argument given in [15], which develops results of [20] and [24].

Proof. We first consider the quantity I_N evaluated for the conditional distribution $P_{XY|AB,cz} = P_{XY|ABCZ}(\cdot, \cdot | \cdot, \cdot, c, z)$, for any fixed c and z . The idea is to use this quantity to bound the trace distance between the conditional distribution $P_{X|acz}$ and its negation, $1 - P_{X|acz}$, which corresponds to the distribution of X if its values are interchanged. If this distance is small, it follows that the distribution $P_{X|acz}$ is roughly uniform.

Let $P_{\bar{X}}$ be the uniform distribution on X . For $a_0 := 0$, $b_0 := 2N - 1$, we have

$$\begin{aligned} I_N(P_{XY|AB,cz}) &= P(X = Y|A = a_0, B = b_0, C = c, Z = z) + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y|A = a, B = b, C = c, Z = z) \\ &\geq D(1 - P_{X|a_0b_0cz}, P_{Y|a_0b_0cz}) + \sum_{\substack{a,b \\ |a-b|=1}} D(P_{X|abcz}, P_{Y|abcz}) \\ &= D(1 - P_{X|a_0cz}, P_{Y|b_0cz}) + \sum_{\substack{a,b \\ |a-b|=1}} D(P_{X|acz}, P_{Y|bcz}) \\ &\geq D(1 - P_{X|a_0cz}, P_{\bar{X}}) \\ &= 2D(P_{X|a_0b_0cz}, P_{\bar{X}}) . \end{aligned} \tag{S7}$$

The first inequality follows from the fact that $D(P_{X|\Omega}, P_{Y|\Omega}) \leq P(X \neq Y|\Omega)$ for any event Ω (see Lemma 2 below). Furthermore, we have used the non-signalling conditions $P_{X|abcz} = P_{X|acz}$ (from (S4)) and $P_{Y|abcz} = P_{Y|bcz}$ (from (S5)), and the triangle inequality for D . By symmetry, this relation holds for all a and b . We hence obtain $D(P_{X|abcz}, P_{\bar{X}}) \leq \frac{1}{2}I_N(P_{XY|AB,cz})$ for all a, b, c and z .

We now take the average over z on both sides of (S7). The left-hand-side gives

$$\begin{aligned} \sum_z P_{Z|abc}(z)I_N(P_{XY|AB,cz}) &= \sum_z P_{Z|c}(z)I_N(P_{XY|AB,cz}) \\ &= \sum_z P_{Z|a_0b_0c}(z)P(X = Y|a_0, b_0, c, z) + \sum_{\substack{a,b \\ |a-b|=1}} \sum_z P_{Z|abc}(z)P(X \neq Y|a, b, c, z) \\ &= P(X = Y|a_0, b_0, c) + \sum_{\substack{a,b \\ |a-b|=1}} P(X \neq Y|a, b, c) \\ &= I_N(P_{XY|AB,c}), \end{aligned} \tag{S8}$$

where we used the non-signalling condition $P_{Z|abc} = P_{Z|c}$ (which is implied by (S4) and (S5)) several times. Furthermore, taking the average on the right-hand-side of (S7) yields $\sum_z P_{Z|abc}(z)D(P_{X|abcz}, P_{\bar{X}}) = D(P_{XZ|abc}, P_{\bar{X}} \times P_{Z|abc})$, so we have

$$2D(P_{XZ|abc}, P_{\bar{X}} \times P_{Z|abc}) \leq I_N(P_{XY|AB,c}) = I_N(P_{XY|AB}), \tag{S9}$$

where the last equality follows from the non-signalling condition (S3).

Inequality (S9) and the relation $D(P_X, Q_X) \leq D(P_{XY}, Q_{XY})$ imply $D(P_{X|abc}, P_{\bar{X}}) \leq \frac{1}{2}I_N(P_{XY|AB})$, and hence

$$\left| P_{X|abc}(x) - \frac{1}{2} \right| \leq \frac{1}{2}I_N(P_{XY|AB}) \tag{S10}$$

for all a, b, c and x . Furthermore, since

$$2D(P_{XZ|abc}, P_{\bar{X}} \times P_{Z|abc}) = \sum_z \left| P_{XZ|abc}(0, z) - \frac{1}{2}P_{Z|abc}(z) \right| + \sum_z \left| P_{XZ|abc}(1, z) - \frac{1}{2}P_{Z|abc}(z) \right|,$$

and both terms on the right-hand-side are equal, using (S9) we have

$$\sum_z \left| P_{XZ|abc}(x, z) - \frac{1}{2}P_{Z|abc}(z) \right| \leq \frac{1}{2}I_N(P_{XY|AB}),$$

for all a, b, c and x . Combining this with (S10) gives

$$\begin{aligned} D(P_{Z|abcx}, P_{Z|abc}) &= \sum_z \left| \frac{1}{2}P_{Z|abcx}(z) - \frac{1}{2}P_{Z|abc}(z) \right| \\ &\leq \sum_z \left| \frac{1}{2}P_{Z|abcx}(z) - P_{X|abc}(x)P_{Z|abcx}(z) \right| + \sum_z \left| P_{X|abc}(x)P_{Z|abcx}(z) - \frac{1}{2}P_{Z|abc}(z) \right| \\ &= \sum_z P_{Z|abcx}(z) \left| \frac{1}{2} - P_{X|abc}(x) \right| + \sum_z \left| P_{XZ|abc}(x, z) - \frac{1}{2}P_{Z|abc}(z) \right| \\ &\leq I_N(P_{XY|AB}). \end{aligned}$$

This establishes the relation (S6). □

Lemma 2. *Let X and Y be random variables jointly distributed according to P_{XY} . The variational distance between the marginal distributions P_X and P_Y is bounded by*

$$D(P_X, P_Y) \leq P(X \neq Y).$$

Proof. Let $P_{XY}^\neq := P_{XY|X \neq Y}$ be the joint distribution of X and Y conditioned on the event that they are not equal. Similarly, define $P_{XY}^\equiv := P_{XY|X=Y}$. We then have

$$P_{XY} = p_\neq P_{XY}^\neq + (1 - p_\neq) P_{XY}^\equiv$$

where $p_{\neq} := P(X \neq Y)$. By linearity, the marginals of these distributions satisfy the same relation, i.e.,

$$P_X = p_{\neq} P_X^{\neq} + (1 - p_{\neq}) P_X^{\bar{}} \quad \text{and} \quad P_Y = p_{\neq} P_Y^{\neq} + (1 - p_{\neq}) P_Y^{\bar{}}.$$

Hence, by the convexity of the variational distance,

$$D(P_X, P_Y) \leq p_{\neq} D(P_X^{\neq}, P_Y^{\neq}) + (1 - p_{\neq}) D(P_X^{\bar{}}, P_Y^{\bar{}}) \leq p_{\neq},$$

where the last inequality follows because the variational distance cannot be larger than one, and $D(P_X^{\bar{}}, P_Y^{\bar{}}) = 0$. \square

PART III OF THE PROOF

In this section we give the proof of the final part of Theorem 1.⁴ We use the setup and assumptions as formulated at the beginning of the Supplementary Methods. In Parts I and II of the proof (see the main text and the previous section) we showed that for all a, b, c and x , the relation $P_{Z|acx} = P_{Z|ac}$ holds for projective quantum measurements compatible with one half of a maximally entangled state (cf. Lemma 1 and recall that for such measurements, the quantity I_N can be made arbitrarily small for sufficiently large N). Part III, explained here, extends this claim to arbitrary states (not necessarily maximally entangled ones) and arbitrary measurements.

The argument proceeds in two steps. The first is to reduce the problem to a situation where the measurement outcome is essentially uniform. Let $(A \rightsquigarrow X, \{E_x^{\bar{a}}\}_x, \mathcal{H}_S)$ be the quantum measurement under consideration (where the input $A = \bar{a}$ is fixed). The idea is that we can always append a second measurement, generating \bar{X} , such that the distribution of the joint output (X, \bar{X}) is flat (to any desired accuracy).

Lemma 3. *Let $\varepsilon > 0$ and let ρ_S be an arbitrary density operator on \mathcal{H}_S . For any measurement on S there exists an additional measurement such that the joint output distribution of (X, \bar{X}) , obtained by applying the two measurements sequentially to ρ_S , has distance ε to a flat distribution.*

Proof idea. It is easy to see that any probability distribution can be turned into an approximately flat one by adding an additional random process that “splits” each probability into sufficiently many smaller events. Furthermore, any such random process can be obtained by an appropriate choice of projective measurement (in a sufficiently large Hilbert space). \square

Let $\{E_{x,\bar{x}}^{\bar{a}}\}_{x,\bar{x}}$ be the set of measurement operators corresponding to the measurement $(A \rightsquigarrow (X, \bar{X}), \{E_{x,\bar{x}}^{\bar{a}}\}_{x,\bar{x}}, \mathcal{H}_S)$ which generates the pair (X, \bar{X}) , and let ρ_S be a pure quantum state compatible with this measurement (see Assumption *QMa*). Next, we introduce projectors $\{\Pi_{x,\bar{x}}\}_{x,\bar{x}}$ and an isometry \mathcal{E} such that $\sigma_{SDR} = \mathcal{E}(\rho_S)$ satisfies

$$\text{tr}_{DR}((\mathbb{1}_S \otimes \Pi_{x,\bar{x}} \otimes \mathbb{1}_R) \sigma_{SDR} (\mathbb{1}_S \otimes \Pi_{x,\bar{x}} \otimes \mathbb{1}_R)) = E_{x,\bar{x}}^{\bar{a}} \rho_S (E_{x,\bar{x}}^{\bar{a}})^{\dagger}.$$

(Note that the isometry can always be defined such that the projectors $\Pi_{x,\bar{x}}$ have rank one.) According to Assumption *QMb* we can append additional quantum measurements $(A' \rightsquigarrow (X', \bar{X}'), \{F_{x,\bar{x}}^a\}_{a,x,\bar{x}}, \mathcal{H}_D)$ (with $F_{x,\bar{x}}^{\bar{a}} = \Pi_{x,\bar{x}}$) and $(B \rightsquigarrow Y, \{G_y^b\}_{b,y}, \mathcal{H}_S \otimes \mathcal{H}_R)$, such that the output statistics are compatible with σ_{SDR} . Furthermore, $(X', \bar{X}') = (X, \bar{X})$ whenever $A' = \bar{a}$. Finally, by Assumption *FR* we can take A' and B to be free choices with $O_A \rightsquigarrow (X', \bar{X}')$, $O_B \rightsquigarrow Y$, and $O_C \rightsquigarrow Z$ spacelike separated (where O_A , O_B , and O_C are the trigger events for A' , B , and C , respectively).

Since the outcomes (X', \bar{X}') of the measurement (for $A' = \bar{a}$) are almost (up to an arbitrarily small distance ε) uniformly distributed, and the state σ_{SDR} is pure, it must be (almost) maximally entangled between the measurement device, \mathcal{H}_D and the remaining systems, $\mathcal{H}_S \otimes \mathcal{H}_R$ (by a suitable choice of the additional measurement, we can always take this to be maximally entangled over an integer number of two-level systems). Furthermore, $\{\Pi_{x,\bar{x}}\}_{x,\bar{x}}$ are orthogonal projectors. Hence, by a suitable choice of the additional measurements producing (X', \bar{X}') and Y , the argument given in Parts I and II of the proof implies that, for any $\varepsilon > 0$ and for all c, x and \bar{x} ,

$$D(P_{Z|A'=\bar{a},cx\bar{x}}, P_{Z|A'=\bar{a},c}) \leq \varepsilon.$$

Since the values of (X, \bar{X}) and (X', \bar{X}') coincide for $A' = A = \bar{a}$ (cf. Assumption *QMb*), we have

$$D(P_{Z|A=\bar{a},cx}, P_{Z|A=\bar{a},c}) \leq \varepsilon.$$

This relation holds for all \bar{a} , and, since ε can be arbitrarily small, establishes the desired Markov chain condition $P_{Z|A=\bar{a},cx} = P_{Z|A=\bar{a},c}$.

⁴ The proof we give here is similar to an argument given by Zurek [27] to derive the Born rule starting from unitarity.

REMARKS ON THE NOTION OF LOCALITY

Here we make some comments about the notion of locality. The main point is to highlight that Bell's notion of locality is similar to, but slightly less general than, the non-signalling nature of the extension (as derived in Part I of the proof).

To quote Bell [2], *locality* is the requirement that "...the result of a measurement on one system [is] unaffected by operations on a distant system with which it has interacted in the past..." Indeed, our non-signalling conditions reflect this requirement and, in our language, the statement that $P_{XYZ|ABC}$ is non-signalling is equivalent to a statement that the model is *local* (see also the discussion in [28]). (We remind the reader that we do not assume the non-signalling conditions, but instead derive them from the free choice assumption.)

In spite of the above quote, Bell's formal definition of locality is slightly more restrictive than these non-signalling conditions. Bell considers extending the theory using hidden variables, here denoted by the variable Z . He requires $P_{XY|ABZ} = P_{X|AZ} \times P_{Y|BZ}$ (see e.g. [13]), which corresponds to assuming not only $P_{X|ABZ} = P_{X|AZ}$ and $P_{Y|ABZ} = P_{Y|BZ}$ (the non-signalling constraints, also called *parameter-independence* in this context), but also $P_{X|ABYZ} = P_{X|ABZ}$ and $P_{Y|ABXZ} = P_{Y|ABZ}$ (also called *outcome-independence*). These additional constraints do not follow from our assumptions and are not used in this work.

A possible reason for the discrepancy is that Bell principally considered extended theories which are deterministic given the hidden variables. In this case, the distinction between Bell's notion of locality and the non-signalling conditions we use is unimportant: if X is deterministic given A and Z , then $P_{X|ABYZ} = P_{X|AZ}$ follows automatically. In fact, the converse also holds: given parameter-independence and outcome-independence a necessary condition for the model to recreate the quantum correlations arising from measurements on a maximally entangled state is that it is deterministic given the hidden variables. To see this, note that for any measurement $A = a$, there is a corresponding measurement $B = b_a$ such that quantum theory predicts identical outcomes. In other words, $P_{X|ab_ayz} = \delta_{x,y}$. The assumptions of parameter-independence and outcome-independence give $P_{X|az} = P_{X|ab_ayz}$, and so $P_{X|az}(x) = \delta_{x,y}$. This implies that X and Y are determined given A and Z .

CANDIDATE EXTENSIONS BASED ON SIMULATIONS OF QUANTUM CORRELATIONS

It has been shown in a number of ways that quantum correlations can be simulated from other resources. For example, all correlations generated by projective measurements on a maximally entangled pair of qubits can be simulated by shared randomness and one bit of classical communication [29], or by shared randomness and a non-local box [30] (a hypothetical device with stronger-than-quantum correlations [31, 32]). Furthermore, these results have been generalized to arbitrary (not necessarily maximally entangled) pure states [33].

Since such simulations recreate quantum correlations, they may appear at first sight to be extensions of quantum theory. We will not provide an exhaustive treatment of all such models, but instead give a short explanation as to why the examples above do not contradict our claim.

First note that the ability to simulate quantum correlations does not imply the ability to predict the outcomes of a genuine quantum experiment. However, when thinking about these simulations in the context of extending quantum theory, the hypothesis is that the components of the simulation really exist and are used to generate outcomes.

The case where communication is needed is analogous to de Broglie-Bohm theory [16, 17] (discussed in the main text). In order that the simulation can work in the case of spacelike separated measurements, the communication bit, Z (which depends on one of the measurement choices, say A), must propagate faster than light. The bit Z is therefore accessible outside the future lightcone of A . According to Assumption *FR*, it must be possible to choose A to be independent of this (now pre-existing) information, which would no longer be the case. Such models therefore contradict Assumption *FR*.

In the model of [30], where a non-local box is used for the simulation, even with full access to this box, there is no better way to predict the measurement outcomes. To see this, note that the output, X , of a measurement specified by a parameter, A , is generated in the simulation by XORing a shared classical value with the output of a non-local box, whose input depends on A . Since the individual outputs of a non-local box are uniform and random the same is true for X . Hence, while the simulation recreates the correct quantum correlations, it does not extend quantum theory in the sense of providing any extra information about future measurement outcomes. It is hence in agreement with Part II of the proof

However, because it recreates the quantum correlations, the simulation provides more information about the outcomes of joint measurements. To see that this is incompatible with quantum theory, one would need to apply Part III of our argument, using a description of how the model evolves under reversible operations. Such a description is not given in the above model and, furthermore, in consistent theories which permit non-local boxes [34] the reversible

dynamics are known to be trivial [35]. They cannot therefore result in a state whose statistics are consistent with those from a quantum evolution, and hence contradict Assumption *QMb*.

SUPPLEMENTARY REFERENCES

- [27] Zurek, W. H. Relative states and the environment: Einselection, enviance, quantum Darwinism, and the existential interpretation. e-print [arXiv:0707.2832](https://arxiv.org/abs/0707.2832) (2007).
- [28] Hall, M. J. W. Comment on ‘Non-realism: deep thought or a soft option?’, by N. Gisin. e-print [arXiv:0909.0015](https://arxiv.org/abs/0909.0015) (2009).
- [29] Toner, B. & Bacon, D. Communication cost of simulating Bell correlations. *Physical Review Letters* **91**, 187904 (2003).
- [30] Cerf, N. J., Gisin, N., Massar, S. & Popescu, S. Simulating maximal quantum entanglement without communication. *Physical Review Letters* **94**, 220403 (2005).
- [31] Cirel’son, B. Some results and problems on quantum Bell-type inequalities. *Hadronic Journal Supplement* **8**, 329–345 (1993).
- [32] Popescu, S. & Rohrlich, D. Quantum nonlocality as an axiom. *Foundations of Physics* **24**, 379–385 (1994).
- [33] Brunner, N., Gisin, N., Popescu, S. & Scarani, V. Simulation of partial entanglement with nonsignaling resources. *Physical Review A* **78**, 052111 (2008).
- [34] Barrett, J. Information processing in generalized probabilistic theories. *Physical Review A* **75**, 032304 (2007).
- [35] Gross, D., Müller, M., Colbeck, R. & Dahlsten, O. C. O. All reversible dynamics in maximally nonlocal theories are trivial. *Physical Review Letters* **104**, 080402 (2010).