

ARTICLE

Received 3 Oct 2012 | Accepted 28 Mar 2013 | Published 14 May 2013

DOI: 10.1038/ncomms2821

An information-theoretic principle implies that any discrete physical theory is classical

Corsin Pfister¹ & Stephanie Wehner^{1,2}

It has been suggested that nature could be discrete in the sense that the underlying state space of a physical system has only a finite number of pure states. Here we present a strong physical argument for the quantum theoretical property that every state space has infinitely many pure states. We propose a simple physical postulate that dictates that the only possible discrete theory is classical theory. More specifically, we postulate that no information gain implies no disturbance or, read in the contrapositive, that disturbance leads to some form of information gain. Furthermore, we show that non-classical discrete theories are still ruled out even if we relax the postulate to hold only approximately in the sense that no information gain only causes a small amount of disturbance. Our postulate also rules out popular generalizations such as the Popescu-Rohrlich-box that allows non-local correlations beyond the limits of quantum theory.

¹Centre for Quantum Technologies, National University of Singapore, 3 Science Drive 2, Singapore 117543, Singapore. ²School of Computing, National University of Singapore, 13 Computing Drive, Singapore 117417, Singapore. Correspondence and requests for materials should be addressed to C.P. (email: mail@corsinpfister.com).

In contrast to classical theory, quantum theory has the remarkable property that the state space of every system has continuously many pure states. These are states that can be seen as states of maximal knowledge: they cannot be prepared by flipping a (possibly biased) coin to decide between two different preparation procedures to be executed, hiding the outcome of the coin flip. Even the qubit, the smallest possible system with no more than two perfectly distinguishable states, has continuously many such states. This non-discreteness of quantum theory contrasts sharply with classical theory, where systems with a finite number of perfectly distinguishable states have the same finite number of pure states. While from a mathematical point of view, this quantum property is satisfactorily explained as a consequence of the mathematical framework of quantum theory, a physical explanation of this phenomenon is less evident.

Indeed, one might conjecture that the actual state space of a physical system really was discrete with only finitely many pure states (Fig. 1)^{1,2}. The fact that experiments have not found a deviation from the continuous nature of the quantum state spaces could then be explained by insufficient measurement precision. A qubit, for example, could be described by a polytope that approximates the continuous spherical shape of the Bloch ball very well, while it actually is a discrete system. Quantum gravitational considerations have led some authors to the idea that indications for the discreteness of spacetime could in turn provide an indication for the discreteness of quantum state spaces^{1,2}. Such considerations might suggest state spaces with an extremely high number of pure states, but as long as the number of pure states is finite, they would differ from quantum state spaces in a fundamental way.

In this work, we present a strong physical counter-argument to the idea that quantum theory could be replaced by a theory with discrete state spaces. This argument is derived from a postulate that claims a very basic principle for measurements. It states that every (pure) measurement can be performed in a way such that the states with a definite outcome (that is, the states with an outcome of probability one) are left invariant. We regard this principle to be a natural property of a theory that describes physical measurements, so we impose it as a postulate. Performing a measurement with a definite outcome does not give any information, while performing a measurement for which the outcome is not known in advance can be seen as a process of gaining information. This allows to regard our postulate as a converse to the well-known fact in quantum theory that information gain causes disturbance³: we postulate that a measurement with no information gain causes no disturbance.

We prove that a non-classical probabilistic theory that satisfies this postulate cannot be discrete. By a discrete system, we mean a system for which the state space has only finitely many pure states. In other words, we show that every theory that satisfies our postulate must either be classical or it must have infinitely many pure states.

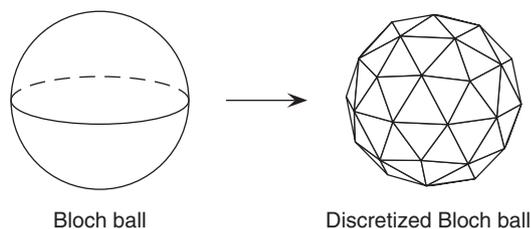


Figure 1 | Illustration of discretized state spaces. One might conjecture that physical state spaces are discrete in the sense that they only have a finite number of pure states. In such discrete state spaces, the pure states are given by the corners of the state space.

Results

The framework. We formulate our result in the abstract state space framework^{4–7}. This framework arises from the idea to consider the largest possible class of physical theories (more precisely, generalized probabilistic theories) which satisfy minimal assumptions, containing classical and quantum theory as special cases. This allows us to study properties of quantum theory, like the non-discreteness of the state space, from an outside perspective. Here, we discuss these minimal assumptions very briefly and refer to Pfister⁸ for a detailed introduction to the abstract state space framework and its mathematical background.

The framework, which relies on four minimal assumptions, is based on the idea that any physical theory admits the notions of states and measurements. Their interpretation is assumed to be given. The first assumption is that the normalized states form a convex subset Ω_A of a real vector space A . The underlying motivation is the idea of probabilistic state preparation: if $\omega, \tau \in \Omega_A$ are states that can each be prepared by a corresponding preparation procedure, then executing the preparation procedures with probability p and $1-p$ should also lead to a state (described by the convex sum $p\omega + (1-p)\tau$), which should therefore be an element of Ω_A as well. The second assumption is that the dimension of the vector space containing the set of states is arbitrarily large but finite. This is a purely technical assumption intended to make the involved mathematics feasible. The third assumption is that the set of states Ω_A is compact. Although there might be some physical motivation for this assumption, we shall be satisfied with considering it as a technical assumption.

Before we discuss the fourth assumption, we make a few comments on the structure of Ω_A . The extreme points of Ω_A are the pure states of the system, the other elements are called mixed states. As Ω_A is a convex and compact subset of a finite-dimensional vector space A , every element of Ω_A is a convex combination of the extreme points of Ω_A (ref. 9). Thus, every state is a convex combination of pure states. As a convex combination is a sum with positive weights that sum up to one, a state can be seen as a probability distribution over pure states. In general, this probability distribution is not unique. In classical theory, however, it is (see the example below). In addition to the normalized states Ω_A , an abstract state space A also contains the subnormalized states $\Omega_A^{\leq 1}$, which are given by all rescalings of the normalized states by factors between zero and one.

The fourth assumption states, roughly speaking, that every mathematically well-defined measurement is regarded as a valid measurement: a measurement is a finite set $\mathcal{M} = \{f_1, \dots, f_n\}$ of functions $f_i: A \rightarrow \mathbb{R}$ that are called effects, each corresponding to an outcome of the measurement. For a state $\omega \in \Omega_A$, the value $f_i(\omega)$ is interpreted to be the probability that the measurement yields the outcome i when the system was in the state ω before the measurement. Thus, one must have $0 \leq f_i(\omega) \leq 1$ for all $\omega \in \Omega_A$. If the measured system was in the state ω with probability p and in the state τ with probability $1-p$, then the probability $pf_i(\omega) + (1-p)f_i(\tau)$ of getting the outcome i has to be identical to $f_i(p\omega + (1-p)\tau)$ as $p\omega + (1-p)\tau$ is regarded to be a state in its own right (in accordance with the first assumption). Skipping a few details, this means that effects are assumed to be linear. Moreover, the effects of a measurement have to sum up to the so-called unit effect $\sum_{i=1}^n f_i = u_A$ for which $u_A(\omega) = 1$ for all $\omega \in \Omega_A$ (as the probability that any outcome occurs has to be one). The fourth assumption is that every set of such linear functionals (effects) is a valid measurement. We denote the set of all effects on an abstract state space by E_A , and we denote measurements (that is, sets of effects that sum up to the unit effect) by calligraphic letters (\mathcal{M} or \mathcal{N} in this paper).

We would like to emphasize that the fourth assumption, which connects the geometry of the states with the geometry of the

effects⁸, is standard but non-trivial and of crucial technical importance for our result. A compelling physical motivation does not seem to be obvious, so it should be regarded as a tentative assumption on the way to a better understanding of quantum theory. Note that as a consequence of this assumption, a theory where the set of states is a quantum state space but where the measurements are restricted to a proper subset of the positive operator valued measures (POVMs) is not part of the framework (c.f. quantum theory in the examples below). In quantum information science, it is always assumed that the full set of POVMs can be performed.

These four assumptions determine the framework of abstract state spaces. This structure is sufficient as long as one is only interested in measurement statistics of one-shot measurements. If one wants to describe several consecutive measurements, one has to introduce measurement transformations. We will discuss this below, but first, we make a few examples.

In the following, we introduce a few examples of theories that can be formulated in the abstract state space framework. (More examples can be found in ref. 8.) While quantum and classical theories are theories of actual physical significance, other theories that we introduce have the role of toy theories, which are helpful to understand the framework. Especially the square and the pentagon, which are instances of polygon models (see below), will serve as useful examples in the illustration of the proof idea of our result.

As a first example, let us have a look at quantum theory. The set of states of a (finite-dimensional) quantum system is given by $\Omega_A = \mathcal{S}(\mathcal{H})$ for some (finite-dimensional) Hilbert space \mathcal{H} , where $\mathcal{S}(\mathcal{H})$ denotes the positive operators on \mathcal{H} with unit trace (the density operators). These operators form a compact convex subset of $A = \text{Herm}(\mathcal{H})$, the vector space of Hermitian operators on \mathcal{H} . Every quantum system has continuously many pure states. The most general description of measurement statistics in quantum theory is given by a POVM, which is a set $\{F_i\}_{i=1}^n$ of positive operators that sum up to the identity-operator I on \mathcal{H} . They give rise to the effects $\rho \mapsto \text{tr}(F_i \rho)$ that sum up to the unit effect u_A given by $\rho \mapsto \text{tr}(I \rho) = 1$ for all $\rho \in \mathcal{S}(\mathcal{H})$. In analogy to our comment above, we emphasize that a theory where the states form a proper subset of a quantum state space but where the measurements are given by not more than POVMs fails to satisfy the fourth assumption of the framework because a reduction of the allowed states requires an extension of the effects.

Another example is classical theory. The states Ω_A of a (finite) classical theory are given by a simplex, that is, by the convex hull of finitely many affinely independent points. (We say that points p_1, \dots, p_n in a real vector space are affinely independent if no point is an affine combination of the other points, that is, if for every p_i , there are no real coefficients $\{\alpha_k\}_{k \neq i}$ with $\sum_{k \neq i} \alpha_k = 1$ such that $\sum_{k \neq i} \alpha_k p_k = p_i$.) Examples of simplices are given by a line segment, a triangle, a tetrahedron, a pentachoron and so on. Every element of a simplex Ω_A is a unique convex combination of the extreme points of Ω_A (Fig. 2). Thus, for a simplex Ω_A , the states are in a one-to-one correspondence with the probability distributions over the pure states, which in the case of a simplex are perfectly distinguishable. This allows to interpret the pure states as classical symbols. In a classical system, there is a generic measurement. For a given state ω , the outcome probabilities for this measurement are precisely the coefficients in the convex sum of the pure states that yield ω .

A more general class of examples is given by what we call discrete theories. We say that Ω_A is a discrete state space if it is the convex hull of finitely many (not necessarily affinely independent) points. As Ω_A is compact, this is equivalent to saying that the theory has only finitely many pure states. Classical theory is an example of a discrete theory, while quantum theory is not.

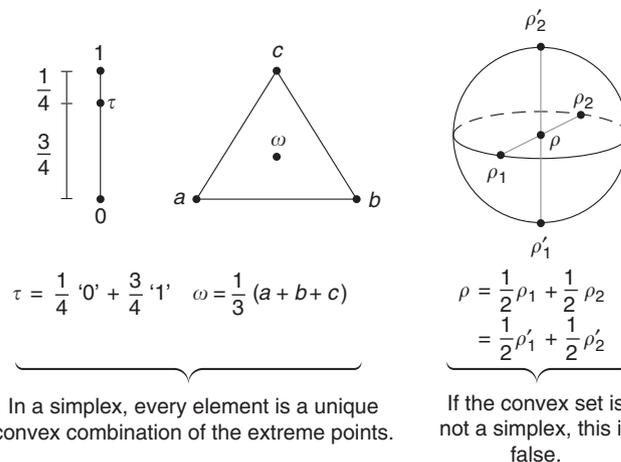


Figure 2 | (Non-)Uniqueness of convex decompositions. A classical system is described by a simplex, which has the property that every point is a unique convex combination of the extreme points. Thus, a state in a classical system corresponds to a unique probability distribution over classical symbols.

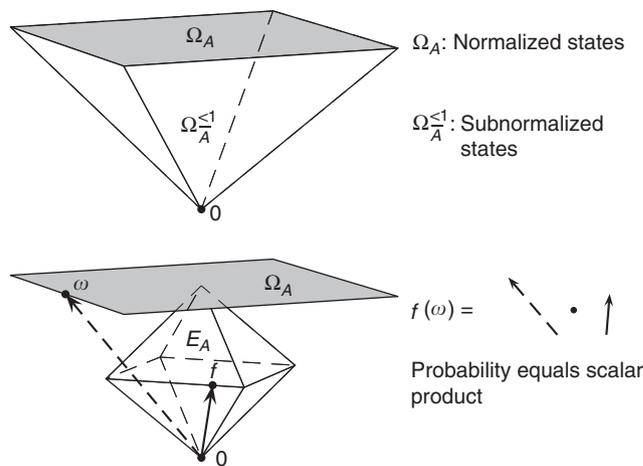


Figure 3 | The square polygon model. The upper part of the figure shows the set of normalized states Ω_A (grey), together with the subnormalized states $\Omega_A^{\leq 1}$ (white ‘pyramid’), which are given by all rescalings of normalized states with factors between zero and one. In the lower part of the figure, the subnormalized states are omitted. Instead, the effects E_A are shown (here they correspond to an octahedron). The reader who is familiar with the mathematics of ordered vector spaces may notice that the effects arise from the structure of the dual cone A^*_+ (more precisely, the effects form an order interval $[0, u_A]$ in A^*)⁸. Here, they are represented as vectors in the same space as the states. To calculate a probability $f(\omega)$, one simply takes the scalar product of the vector ω and the vector representing f .

Very illustrative examples are given by the polygon models¹⁰: these are abstract state spaces where Ω_A is a regular polygon, so they are special kinds of discrete theories. As the whole situation can be drawn in only three dimensions, the polygon models provide examples for which we can give a picture (Fig. 3). To see the interplay of states and effects in such a low-dimensional example, it is useful to represent effects as vectors in the same space as the states¹⁰. To evaluate an effect at some state, one simply takes the scalar product of the state and the vector representing the effect. In the Methods section below, the square and the pentagon will be the central examples in the illustration of the proof idea.

So far, we have discussed the core structure of abstract state spaces: states and effects. They only allow for the description of one-shot measurement statistics. If one wants to describe the statistics of several consecutive measurements, then one has to specify what happens to the state of the system when a measurement is performed (otherwise, the statistics of the subsequent measurement cannot be described). In other words, one has to specify a rule for post-measurement states. The structure of an abstract state space, however, does not provide such a rule and leaves open the question of how to specify post-measurement states.

We deal with this question and consider some extra structure on abstract state spaces that provides a rule for post-measurement states. We describe the transition from the initial state of the system (before the measurement) to the post-measurement state by what we call a measurement transformation. Such transformations have been considered, for example, in refs 11–13. We go one step further. Our result makes a statement about the existence of measurement transformations in abstract state spaces that satisfy a certain postulate.

As we have just mentioned above, the general idea is that a measurement transformation specifies a rule for how post-measurement states are assigned. However, in a physical theory, how such a rule looks like depends on the particular situation that one wants to describe. To be more specific, we can think of at least three such situations (we will make quantum examples below), which correspond to the case where (a) the observer finds out the outcome of the measurement and describes the state of the system after the measurement conditioned on that outcome; (b) the observer describes the system after the measurement by a subnormalized state for the hypothetical case that a particular outcome occurred, incorporating the probability of that outcome into the post-measurement state; and (c) the observer does not find out the outcome of the measurement and describes the state of the system after the measurement, knowing only that the measurement has been performed. A physical theory has to allow for a mathematical description for all of these cases. Each of the three situations can be described by a particular kind of map. To understand the difference between them, it is helpful to see how these maps look like for the particular case of quantum theory. There, if the measurement is a projective measurement $\mathcal{M} = \{P_i\}_{i=1}^n$, the maps are given by Lüders projections¹⁴ (the literature is ambiguous about which of the three maps is called a Lüders projection, but as they are very closely related, this usually does not lead to problems). The situations (a), (b) and (c) above are described by the following maps: in situation (a), if the outcome associated with projector P_k is measured, then the state is transformed as

$$\rho \mapsto \frac{P_k \rho P_k}{\text{tr}(P_k \rho)}. \quad (1)$$

In situation (b), considering the outcome associated with projector P_k , the state transforms into a subnormalized state as

$$\rho \mapsto P_k \rho P_k. \quad (2)$$

In situation (c), if the outcome of the measurement is unknown, the state is transformed as

$$\rho \mapsto \sum_{i=1}^n \text{tr}(P_i \rho) \frac{P_i \rho P_i}{\text{tr}(P_i \rho)} = \sum_{i=1}^n P_i \rho P_i. \quad (3)$$

Most introductory textbooks on quantum theory only discuss situation (a). Note that (a) is not a linear map. By the definition that we will make below, it should not be called a transformation.

The maps (b) and (c) are linear. The map (b) describes what Lüders calls a ‘measurement followed by selection’, whereas the map (c) describes what he calls a ‘measurement followed by aggregation’¹⁴.

The preceding discussion allows us to understand what we mean by a measurement transformation. By a measurement transformation, we mean a map of type (b). Note that such a map leads to subnormalized post-measurement states rather than normalized ones. The norm of the post-measurement state (the trace-norm in the quantum case, $\text{tr}(P_k \rho P_k) = \text{tr}(P_k \rho)$) is equal to the probability that the outcome occurs (which is what we mean by ‘the probability of that outcome is incorporated into the state’).

Choosing maps of type (b) (rather than maps of type (a) or (c)) as the subject matter is not a relevant restriction as the three types of maps are so closely related that insights into one of these maps translate into insights into the other maps as well. In particular, from the map of type (b), one can construct the map of type (a) by rescaling the images with the inverse probability and the map of type (c) by summing up over all outcomes.

With the above motivation in mind, we now proceed to the task of formally defining what we mean by a measurement transformation on an abstract state space. A transformation T on an abstract state space A is a linear map $T: A \rightarrow A$ such that $T(\Omega_A) \subseteq \Omega_A^{\leq 1}$. The motivation for the linearity of transformations is similar to the motivation for the linearity of effects. The linearity expresses a compatibility condition for probabilistically prepared states: if the system is in a state ω with probability p and in a state τ with probability $1-p$ before the transformation, then the transformed state $pT(\omega) + (1-p)T(\tau)$ has to coincide with $T(p\omega + (1-p)\tau)$ as $p\omega + (1-p)\tau$ is regarded as a state in its own right. (A more rigorous argument would require $pf(T(\omega)) + (1-p)f(T(\tau)) = f(T(p\omega + (1-p)\tau))$ for all effects f , which eventually boils down to what we have just required.) A measurement transformation has to satisfy one more condition. As we have explained above, a measurement transformation is associated with a particular outcome, or more precisely, with a particular effect. If T is a measurement transformation for an effect f , then we require that the norm $u_A(T(\omega))$ of the transformed state is equal to the probability $f(\omega)$ for measuring the outcome associated with f . In short, we require

$$u_A \circ T = f. \quad (4)$$

In quantum theory, where u_A is given by the trace, this property is satisfied for projective measurements as the Lüders projection gives $\text{tr}(P\rho P) = \text{tr}(P^2\rho) = \text{tr}(P\rho)$.

We will only consider measurement transformations for a special class of effects that we call pure effects. We say that an effect $f \in E_A$ is pure if it is an extreme point of the (convex) set of effects E_A , and we say that a measurement $\mathcal{M} = \{f_1, \dots, f_n\}$ is pure if every effect f_1, \dots, f_n is pure. It turns out that in the case of quantum theory, an effect $F \mapsto \text{tr}(F\rho)$ of a POVM element F is pure if and only if F is a projector⁸. Thus, we only consider measurement transformations for a class of effects that, in the case of quantum theory, reduces to projectors. For this class, the measurement transformations are given by Lüders projections. The fact that we will restrict our considerations to pure effects is not a restriction of the validity of our result. Quite the contrary, this makes our result stronger. As we will see below, our postulate claims a property of measurement transformations for pure effects rather than claiming this property for all effects. This results in a weaker postulate, so every implication derived from this postulate leads to a stronger result. As we will see later, we will restrict the claim of the postulate to an even smaller subclass of effects (see the Methods section and the Supplementary Note 1 for further details).

In a nutshell, a measurement transformation for a pure effect f is a linear map $T: A \rightarrow A$ with $T(\Omega_A) \subseteq \Omega_A^{\leq 1}$ and $u_A \circ T = f$.

The postulate. Before we can formulate our result, we first state our postulate. For a mathematically precise formulation, we refer to the Methods section and the Supplementary Note 1 of this article.

The postulate reads: Every pure measurement can be performed in a way such that the states for which it yields a certain outcome (that is, the states with an outcome of probability one) are left invariant. In more illustrative terms, this can be rephrased by saying that no information gain implies no disturbance.

In more technical terms, the postulate states that for every pure effect $f \in E_A$, there exists an associated measurement transformation T with $u_A \circ T = f$ such that for all states $\omega \in \Omega_A$ with $f(\omega) = 1$, we have that $T(\omega) = \omega$. The existence of such a measurement transformation T is what is meant by saying that there exists a way to perform the measurement. Furthermore, note that without looking at the definition of a measurement transformation, saying that ‘there exists a way to perform the measurement’ may appear trivial by itself. After all, doing nothing and outputting the measurement outcome (associated with) f preserves ω and yields f with probability 1. This case is ruled out by the definition of a measurement transformation. More precisely, note that T must be such that $(u_A \circ T)(\omega') = f(\omega')$ for all states $\omega' \in \Omega_A$. That is, it yields the correct probabilities for any state that we wish to measure. It is interesting to note that the actual proof of our main result only needs an even weaker but rather technical requirement (see the Methods section).

To see the link to information gain, note that the Shannon information content (see, for example, ref. 15) $-\log f(\omega)$ is zero for any outcome of an experiment that occurs with certainty. As such, $f(\omega) = 1$ is equivalent to stating that no information gain occurs. The demand that $T(\omega) = \omega$ says that the state is unchanged, that is, no disturbance has occurred.

Quantum theory and classical theory satisfy this postulate. In quantum theory, for example, if a system is in a state ρ such that a projective measurement $\{P_i\}_{i=1}^n$ has some outcome k with probability $\text{tr}(P_k \rho) = 1$, then the transformation $\rho \mapsto P_k \rho P_k$ leaves the state invariant. Quantum theory even satisfies the postulate in a much stronger form in the sense that little information gain also causes only little disturbance. This can be seen from a special case of the gentle measurement lemma^{16,17}. It states that if measuring an outcome associated with a projector F has probability $\text{tr}(F \rho) \geq 1 - \epsilon$, then measuring that outcome disturbs the state by no more than $\|\rho - F \rho F\|_1 \leq \sqrt{8\epsilon}$. Setting $\epsilon = 0$, this reduces to our postulate. However, we emphasize that our postulate is much weaker than postulating the gentle measurement lemma. We also note that our postulate does not make any assumptions about locality, that is, it does not make a statement about whether verification measurements of bipartite states can be implemented on local quantum systems or locally disturb the state as has been considered in Popescu and Vaidman¹⁸.

Even though the statement of the postulate is very concise, it may appear unsatisfying as it involves the abstract concept of a state, which is something that one cannot observe directly. However, it can be reformulated in purely operational terms, referring only to directly observable objects, namely measurement statistics. Such a reformulation is possible because two states can be regarded as being identical if and only if they induce the same measurement statistics for every measurement (in more mathematical terms, a state ω is an equivalence class under the relation $\omega \sim \omega' \Leftrightarrow (f(\omega) = f(\omega'))$ for all $f \in E_A$)¹⁹. Hence, instead of making statements about states, one can make statements about the statistics of all potential measurements. Figure 4 illustrates the idea of this reformulation.

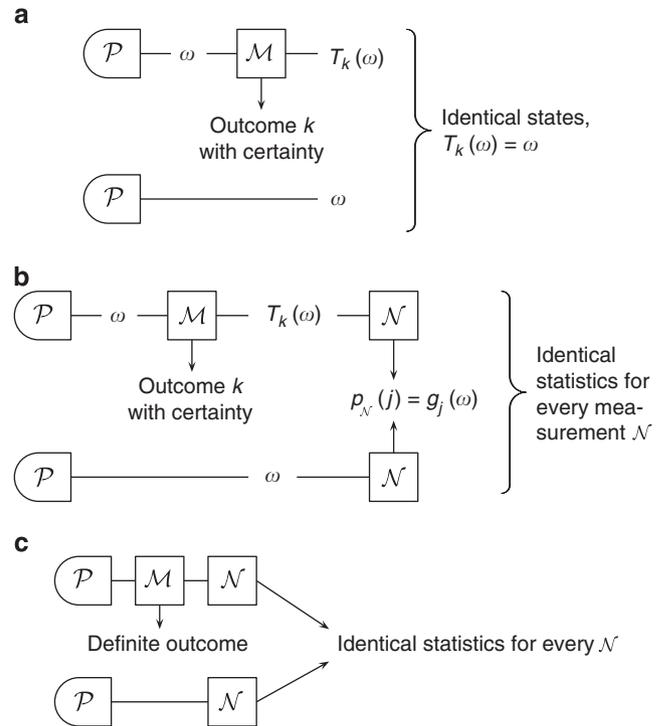


Figure 4 | A reformulation of the postulate in purely operational terms.

Instead of referring to initial and post-measurement states, the reformulated version states that a measurement with a definite outcome does not influence the statistics of any subsequent measurement, so it only refers to directly observable quantities. This reformulation can be understood as follows: consider a preparation \mathcal{P} that outputs an initial state $\omega \in \Omega_A$ and a measurement $\mathcal{M} = \{f_1, \dots, f_n\}$ such that $f_k(\omega) = 1$ for some k . According to the postulate, the state of the system after the two experiments shown in (a) are identical. Thus, if the two experiments are followed by any measurement, say $\mathcal{N} = \{g_1, \dots, g_l\}$, then the statistics of the \mathcal{N} -measurement coincide (see part (b) of the figure). The \mathcal{N} -statistics coincide for every measurement \mathcal{N} . This is equivalent to saying that the states before the \mathcal{N} -measurement (that is, $T_k(\omega)$ and ω) are identical. Thus, we do not need to refer to states and can reformulate the postulate as: if a measurement has a definite outcome, then performing this measurement does not influence the statistics of any subsequent measurement. This is shown diagrammatically in part (c) of the figure.

Main findings. In terms of the postulate, our result can now be stated as follows: an abstract state space that satisfies the postulate is either non-discrete (that is, it has infinitely many pure states) or it is classical.

This means that if a physical system is described by an abstract state space where the set of states Ω_A is a polytope that is not a simplex (that is, if it is a discrete non-classical system), then it violates our postulate.

Furthermore, our result is robust in the sense that discrete non-classical theories are ruled out even if the postulate is weakened to an approximate version. To formulate this approximate version of the result, we assume that A is equipped with a norm $\|\cdot\|_A$. This induces a distance function $\text{dist}(\omega, \omega') := \|\omega - \omega'\|_A$ on A . We prove that for every discrete non-classical theory, equipped with some norm $\|\cdot\|_A$, there is a positive number $\epsilon > 0$ such that the implication $f(\omega) = 1 \Rightarrow \|T(\omega) - \omega\|_A \leq \epsilon$ (where T is the measurement transformation for f) cannot be satisfied for every pure effect $f \in E_A$. We prove this approximate case, which is a stronger version of the result, in the Supplementary Note 2.

Discussion

Our simple postulate rules out discrete non-classical theories, while classical and quantum theory satisfy the postulate. Read in the contrapositive, our postulate says that disturbance implies information gain. Any theory that does not satisfy our postulate thus allows for disturbance without a corresponding ability of information gain. Note that even in a theory that *a priori* only defines transformations T , one can define effects as $u_A \circ T$.

We also note that our postulate rules out several alternatives to quantum theory, most notably the famous Popescu–Rohrlich-box (PR-box)^{20–22} that allows a violation of the CHSH inequality²³ far beyond the limits of quantum theory. More specifically, the PR-box achieves the algebraically maximal violation of the CHSH inequality, while still respecting the law that no information can travel faster than light. This is in spirit similar to other approaches such as information causality²⁴, communication complexity assumptions²⁵, the assumption of local quantum mechanics²⁶ or the uncertainty principle²⁷. We emphasize, however, that whereas this is a nice byproduct of our result, our real aim lies in the study of local physical systems with the goal to identify just one postulate that sheds light on the simple question whether the state space should be discrete or continuous. It is very satisfying that this question can be understood by introducing just a single postulate.

One may wonder whether our postulate does in fact rule out all theories but classical and quantum mechanics. To answer this question, let us first be more precise about what we mean by ‘a theory is (not) ruled out by the postulate’. We mentioned in the preceding section that for general abstract state spaces, measurement transformations are not specified, so we cannot make statements saying that the (unique) measurement transformations do (not) satisfy our postulate. Instead, we can discuss the following well-defined question: given an abstract state space, is it true that for every pure effect, there exists a measurement transformation that satisfies our postulate? If this is the case, then we say that the theory can satisfy the postulate, or that it is not ruled out by the postulate. If this is not true, then we say that the theory cannot satisfy the postulate, or that it is ruled out by the postulate.

This is the precise meaning of our statement that ‘discrete non-classical theories are ruled out by the postulate’. Using this terminology, we can identify a class of theories that, in addition to classical and quantum theory, is not ruled out by the postulate: the strictly convex theories can satisfy our postulate. These are theories where the set of normalized states is strictly convex, that is, the boundary contains no line segment. There are more theories that can satisfy the postulate, but we do not know a concise classification. For example, a state space Ω_A formed like a piece of pizza is ruled out by the postulate, while a state space formed like an ice cream cone is not. Figure 5 gives an overview.

In the recent past, there have been several attempts to derive (finite-dimensional) quantum theory within a framework of probabilistic theories^{12,28,29}. The idea is the following. One starts with a very general framework of probabilistic theories (like the abstract state space formalism). Then, one imposes a few physical postulates (our postulate can be seen as one such postulate). If one manages to show that all theories in this framework other than quantum theory are ruled out by these physical postulates, then this can be seen as a physical derivation of quantum theory. As our postulate rules out quite a large fraction of all possible abstract state spaces already (Fig. 5), it seems promising that adding just a few more postulates might be sufficient to rule out all theories except for quantum theory.

However, we do not make such an attempt and focus on one particular aspect only, introducing only one postulate. What

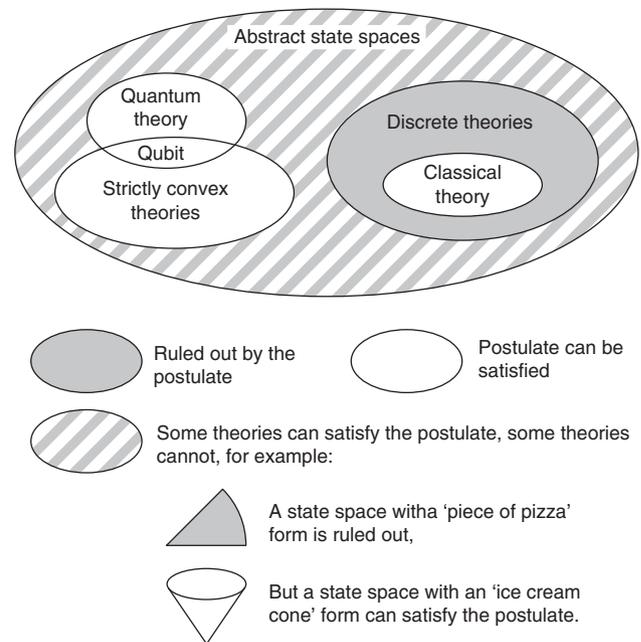


Figure 5 | An overview over the abstract state spaces ruled out by the postulate.

makes our postulate special is that its nature is very different from the postulates that have been considered in this context so far. Many approaches focus on the aspect of non-locality, introducing rules for how physical systems are combined to form bi- or multipartite systems. In contrast, our approach deals with local state spaces only, making a statement about post-measurement states. Within probabilistic theories, this aspect has gained less attention in the literature so far. The fact that, within the framework of abstract state spaces, we introduce just one postulate (instead of a set of postulates) helps us to understand its influence on one particular aspect of physical theories.

One might argue that an experimental proof of the non-discreteness of physical state spaces needs infinite measurement precision as the verification of the postulate that $T(\omega) = \omega$ (strict equality) requires the verification that ω and $T(\omega)$ give rise to the same measurement statistics (to arbitrary precision). Hence, our result is experimentally less accessible than other no-go theorems (for example, the Bell Inequality, where it is sufficient to verify the violation of a single statistical inequality). There is a partial reply to this objection. As we have mentioned before, there is an approximate version of our result. It states that for a given polytope P , there is a positive number $\epsilon_P > 0$ such that the postulate can be weakened to the following form (without changing the validity of the result): if a measurement on a state has an outcome with probability one, then performing the measurement does not change the state of the system by more than ϵ_P (for details, see the Supplementary Note 2). Thus, even if one weakens the postulate to allow for an ϵ_P -disturbance of the state, it still rules out the polytope P . This is a stronger form of the result. It states that in order to rule out a given polytope experimentally, only finite measurement precision is needed (quantified by ϵ_P). However, the allowed disturbance ϵ_P depends on the polytope P , so in order to rule out all polytopes experimentally, infinite measurement precision is needed because for every measurement error, there could be a polytopical theory for the measured system for which the allowed disturbance ϵ_P is too small to be tested.

Methods

Overview. In this section, we sketch the idea of the proof of our main result. This will lead to geometric pictures that illustrate the incompatibility of non-classical discrete state spaces with our postulate (Fig. 6 and Fig. 7). For the full version of the proof and for a proof of the approximate version of our result, see the Supplementary Notes 1 and 2 of this article, respectively.

Here, we aim for a geometric understanding of the proof. It is mainly based on a lemma that establishes geometric criteria for a set of states Ω_A that is compatible with our postulate. To illustrate this lemma, we provide two very basic examples that violate these criteria: the square and the pentagon (Fig. 6). For these two examples, it is easy to see geometrically why they cannot satisfy our postulate (as we will illustrate in Fig. 7). Finally, we describe roughly how we prove that every polytope Ω_A that satisfies the conditions of the lemma is a simplex (which is our main result).

Before we sketch the proof of the main result, it is useful to define in a bit more detail what an abstract state space is. For detailed definitions of the framework see the Supplementary Note 1 of this article, for a detailed motivation of the framework with detailed examples see Chapter 3 in Pfister⁸.

The formal setup. As illustrated in Fig. 8, an abstract state space is fully specified by a tuple (A, A_+, u_A) , where A is a real finite-dimensional vector space, A_+ is a cone in A and u_A is a linear functional on A (called the unit effect). This linear functional is required to be strictly positive on the cone A_+ (that is, $u_A(\omega) > 0$ for all $\omega \in A_+ \setminus \{0\}$). The tuple (A, A_+, u_A) gives rise to the normalized states Ω_A and the subnormalized states $\Omega_A^{\leq 1}$ in the following way (c.f. Fig. 8):

$$\Omega_A := \{\omega \in A_+ \mid u_A(\omega) = 1\}, \tag{5}$$

$$\Omega_A^{\leq 1} := \{\omega \in A_+ \mid u_A(\omega) \leq 1\}. \tag{6}$$

The set E_A of effects on A is given by the linear functionals that take values between zero and one on the states Ω_A , that is

$$E_A := \{f \in A^* \mid 0 \leq f(\omega) \leq 1 \quad \forall \omega \in \Omega_A\}, \tag{7}$$

where A^* is the dual space of A . A measurement is given by a finite set of effects $\mathcal{M} = \{f_1, \dots, f_n\} \subseteq E_A$ such that the effects sum up to the unit effect u_A , that is, $\sum_{i=1}^n f_i = u_A$. Recall that if the system is in the state $\omega \in \Omega_A$ before the measurement described by $\mathcal{M} = \{f_1, \dots, f_n\}$, then the probability for outcome k is given by $f_k(\omega)$.

As we have mentioned earlier, we restrict ourselves to pure effects when we deal with post-measurement states (that is, with measurement transformations). The pure effects are the extreme points of E_A . A pure effect $f \in E_A$ has the property that the set of states ω that have probability $f(\omega) = 1$ is a face of Ω_A ⁸. A face of Ω_A is a convex subset $F \subseteq \Omega_A$ with the property that every line segment whose endpoints are contained in F must be fully contained in F , that is, a face is some kind of ‘extreme subset’. For a pure effect f , this allows us to define the certain face F_f of f by

$$F_f := \{\omega \in \Omega_A \mid f(\omega) = 1\}. \tag{8}$$

Analogously, the set of states ω that have probability $f(\omega) = 0$ is a face of Ω_A as well⁸. We call it the impossible face of f and define it by

$$\bar{F}_f := \{\omega \in \Omega_A \mid f(\omega) = 0\}. \tag{9}$$

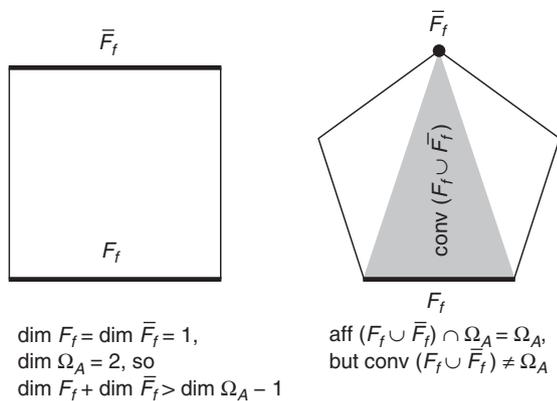


Figure 6 | Violation of the conditions stated in the lemma. The square and the pentagon serve as very basic examples of abstract state spaces that violate the conditions stated in the lemma. The square violates condition (a), while the pentagon violates (b).

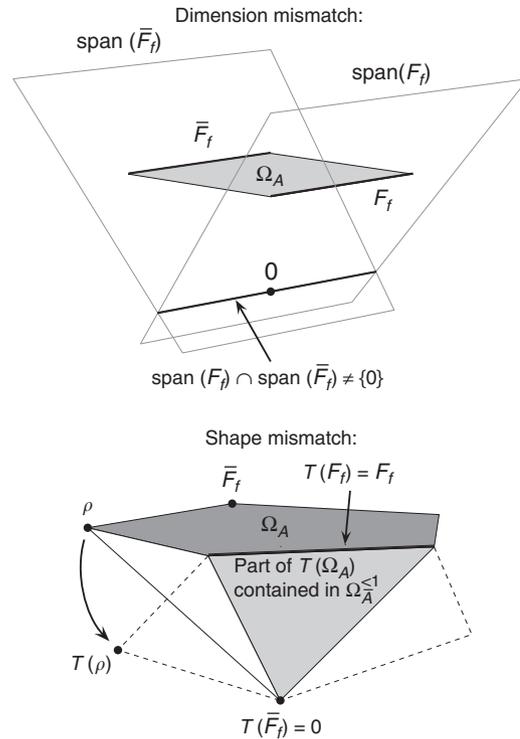


Figure 7 | Consequences of the violation of conditions (a) or (b). This figure illustrates geometrically why the square and the pentagon violate our postulate. Intuitively, all non-classical discrete state spaces exhibit either a dimension or a shape mismatch.

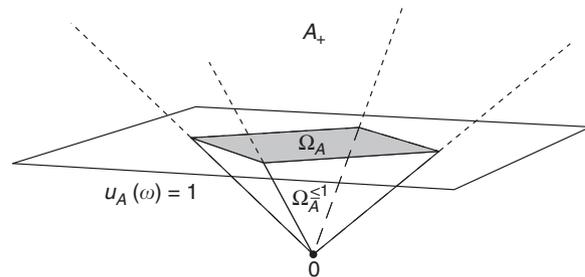


Figure 8 | Visualization of the state cone. The states of any normalization are given by a cone A_+ in the real vector space A . The linear functional u_A gives the normalization of a state, so the intersection of A_+ with the plane described by $u_A(\omega) = 1$ gives the normalized states, while the subnormalized states $\Omega_A^{\leq 1}$ are those elements of A_+ where u_A takes values between 0 and 1.

The notion of the certain face and the impossible face of an effect is central in our proof.

A transformation on an abstract state space is a linear map $T: A \rightarrow A$ that is positive (that is, $T(A_+) \subseteq A_+$) and does not increase the norm of the states, that is, $u_A(T(\omega)) \leq u_A(\omega)$ for all $\omega \in A_+$. Equivalently, a transformation is a linear map $T: A \rightarrow A$ with $T(\Omega_A) \subseteq \Omega_A^{\leq 1}$. Recall that we describe the state change due to a measurement by introducing measurement transformations. If a measurement yields an outcome associated to a pure effect $f \in E_A$, then the transformation of the state is described by $\omega \rightarrow T(\omega)$, where T is the measurement transformation for f . As mentioned, we require that T is a transformation that satisfies $u_A \circ T = f$.

With these definitions at hand, we can formulate our postulate as follows: for every pure effect $f \in E_A$, there is a transformation $T: A \rightarrow A$ such that $f = u_A \circ T$ and $T(\omega) = \omega$ for every $\omega \in F_f$.

Note that we only postulate the existence of a measurement transformation for f that satisfies our postulate. For the actual proof, we will require an even weaker

condition. We will not require the existence of such a measurement transformation for every pure effect but only for pure effects for which the certain face F_f is what we call a minus-face of Ω_A . This is a face that is exactly one dimension smaller than Ω_A . This weakening of the postulate is particularly useful for the proof of the approximate version of our result.

Basic idea of the proof. To derive the result, we first prove a lemma that establishes geometric criteria that a set of states Ω_A has to satisfy to be compatible with our postulate. Given a pure effect $f \in E_A$, the lemma tells us geometric criteria for the certain face F_f and the impossible face \bar{F}_f of f , which are necessary for the existence of a measurement transformation satisfying our postulate.

Let (A, A_+, u_A) be an abstract state space and let $f \in E_A$ be a pure effect. If there exists a transformation $T: A \rightarrow A$ such that $u_A \circ T = f$ and $T(\omega) = \omega$ for every $\omega \in F_f$, then, (a) $\dim F_f + \dim \bar{F}_f \leq \dim \Omega_A - 1$, and (b) $\text{aff}(F_f \cup \bar{F}_f) \cap \Omega_A = \text{conv}(F_f \cup \bar{F}_f)$, where $\text{aff}(\cdot)$ and $\text{conv}(\cdot)$ denote the affine hull and the convex hull, respectively (the reader unfamiliar with these two notions is referred to the Supplementary Note 1).

To get a geometric idea for the two conditions (a) and (b), it is useful to consider abstract state spaces that violate these conditions. The two simplest examples we can think of are the square and the pentagon, depicted in Fig. 6.

To see why the conditions (a) and (b) are necessary for the existence of a transformation compatible with our postulate, we now examine what goes wrong in the case where one of the conditions is violated. If condition (a) is violated, a contradiction occurs that we call a dimension mismatch. If (b) is violated, then we say that a shape mismatch occurs. Again, the square and the pentagon serve as good examples for a geometric illustration.

We first look at the dimension mismatch. If condition (a) is violated (that is, $\dim F_f + \dim \bar{F}_f > \dim \Omega_A - 1$), then there is no linear map T such that

$$u_A \circ T = f, \quad (10)$$

$$T(\omega) = \omega \quad \text{for every } \omega \in F_f \quad (\text{postulate}). \quad (11)$$

In particular, there is no transformation with these two properties. To see this, there are two things to notice.

First, equation (10) implies that $u_A(T(\omega)) = f(\omega) = 0$ for all $\omega \in \bar{F}_f$ (c.f. the definition (9) of \bar{F}_f). As the zero-vector $\omega = 0$ is the only state (that is, the only element of $\Omega_A^{\leq 1}$) for which $f(\omega) = 0$, it follows that the whole impossible face \bar{F}_f has to be mapped to the zero-vector. By the linearity of T , this implies that the restriction $T|_{\text{span}(\bar{F}_f)}$ of T to $\text{span}(\bar{F}_f)$ is the zero-operator on $\text{span}(\bar{F}_f)$:

$$T|_{\text{span}(\bar{F}_f)} = 0|_{\text{span}(\bar{F}_f)}. \quad (12)$$

Second, the postulate (11) and the linearity of T imply that the restriction $T|_{\text{span}(F_f)}$ of T to $\text{span}(F_f)$ is the identity-operator on $\text{span}(F_f)$:

$$T|_{\text{span}(F_f)} = I|_{\text{span}(F_f)}. \quad (13)$$

However, in the case where $\dim F_f + \dim \bar{F}_f > \dim \Omega_A - 1$, equations (12) and (13) lead to a contradiction. In this case, the intersection $\text{span}(\bar{F}_f) \cap \text{span}(F_f)$ is a subspace that is at least one-dimensional (Fig. 7). Equations (12) and (13) imply that on this subspace, T has to be the zero-operator and the identity-operator simultaneously, which could only be satisfied if the subspace would be $\{0\}$.

Now we look at the shape mismatch. If condition (b) is violated (that is, $\text{conv}(F_f \cup \bar{F}_f) \neq \text{aff}(F_f \cup \bar{F}_f) \cap \Omega_A$), then for every linear map that satisfies equations (10) and (11), there is a state ρ such that $T(\rho) \notin \Omega_A^{\leq 1}$ (that is, $T(\rho)$ is not a state). Therefore, such a T cannot be a transformation. To see this geometrically, it is useful to consider the pentagon for a particular choice of the effect f where the certain face F_f is an edge of the pentagon (Fig. 7). Equation (10) implies that the impossible face \bar{F}_f is mapped to the zero-vector, while equation (11) means that the certain face F_f is left invariant. In the case of the pentagon illustrated in Fig. 7, there is precisely one linear map T with these two properties. It maps the normalized states Ω_A (dark grey surface in the figure) to a set in the vector space (dashed lines) which is not contained in $\Omega_A^{\leq 1}$ (the truncated cone between 0 and Ω_A). In particular, there is a ρ such that $T(\rho) \notin \Omega_A$. If one compares Fig. 7 with Fig. 6, then one can see that the part of Ω_A that is mapped to a subset of $\Omega_A^{\leq 1}$ (light grey face in Fig. 7) is precisely given by $\text{conv}(F_f \cup \bar{F}_f)$ (grey part in Fig. 6). However, the part of Ω_A that is mapped outside of Ω_A is given by $(\text{aff}(F_f \cup \bar{F}_f) \cap \Omega_A) \setminus \text{conv}(F_f \cup \bar{F}_f)$ (the white part in Fig. 6). This observation generalizes to statement (b) of the Lemma: if (a) is satisfied, then $T(\Omega_A)$ is contained in $\Omega_A^{\leq 1}$ if and only if $\text{aff}(F_f \cup \bar{F}_f) \cap \Omega_A = \text{conv}(F_f \cup \bar{F}_f)$.

These two examples illustrate all that can go wrong for discrete theories. We show that for every discrete theory (that is, for every theory where Ω_A is a polytope), either condition (10) or (11) is violated (so either a dimension mismatch or a shape mismatch occurs), except for the case where Ω_A is a simplex (that is, for classical theories). To show this, we proceed as follows.

We consider an abstract state space (A, A_+, u_A) where Ω_A is a polytope. We assume that for every pure effect $f \in E_A$ for which the certain face F_f is a minus-face of Ω_A , there is a measurement transformation satisfying the postulate (11).

In a first step, we show (using the lemma) that every polytope Ω_A that is compatible with our postulate has a property that we call being uniformly pyramidal. This means that for every minus-face F of Ω_A , it holds that there is a point $a_F \in \Omega_A$ such

that $\Omega_A = \text{conv}(F \cup \{a_F\})$ (see the Supplementary Note 1 for more intuition). In a second step, we show that every uniformly pyramidal polytope Ω_A is a simplex. This shows that every discrete theory satisfying our postulate has to be classical.

References

- Buniy, R. V., Hsu, S. D. H. & Zee, A. Is Hilbert space discrete? *Phys. Lett. B* **630**, 68–72 (2005).
- Buniy, R. V., Hsu, S. D. H. & Zee, A. Discreteness and the origin of probability in quantum mechanics. *Phys. Lett. B* **640**, 219–223 (2006).
- Fuchs, C. A. & Peres, A. Quantum-state disturbance versus information gain: Uncertainty relations for quantum information. *Phys. Rev. A* **53**, 2038–2045 (1996).
- Barnum, H. & Wilce, A. Ordered linear spaces and categories as frameworks for information-processing characterizations of quantum and classical theory. Preprint at <http://arXiv.org/abs/0908.2354> (2009).
- Barnum, H., Barrett, J., Leifer, M. & Wilce, A. Teleportation in general probabilistic theories. Preprint at <http://arXiv.org/abs/0805.3553> (2008).
- Barnum, H., Gaebler, C. P. & Wilce, A. Ensemble steering, weak self-duality, and the structure of probabilistic theories. Preprint at <http://arXiv.org/abs/0912.5532> (2009).
- Barnum, H. & Wilce, A. Information processing in convex operational theories. *Electron. Notes Theor. Comput. Sci.* **270**, 3–15 (2011).
- Pfister, C. One simple postulate implies that every polytopic state space is classical. Preprint at <http://arXiv.org/abs/1203.5622> (2012).
- Minkowski, H. *Gesammelte Abhandlungen* Vol. 2 (Teubner, 1911).
- Janotta, P., Gogolin, C., Barrett, J. & Brunner, N. Limits on nonlocal correlations from the structure of the local state space. *N. J. Phys.* **13**, 063024 (2011).
- Davies, E. B. & Lewis, J. T. An operational approach to quantum probability. *Comm. Math. Phys.* **17**, 239–260 (1970).
- Hardy, L. Quantum theory from five reasonable axioms. Preprint at <http://arXiv.org/abs/quant-ph/0101012> (2001).
- Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **75**, 032304 (2007).
- Lüders, G. Über die Zustandsänderung durch den Meßprozess. *Ann. Phys.* **443**, 322–328 (1950).
- MacKay, D. J. C. *Information Theory, Inference, and Learning Algorithms* (Cambridge University Press, 2003).
- Winter, A. Coding theorem and strong converse for quantum channels. *IEEE Trans. Inform. Theory* **45**, 2481–2485 (1999).
- Winter, A. The capacity of the quantum multiple-access channel. *IEEE Trans. Inform. Theory* **47**, 3059–3065 (2001).
- Popescu, S. & Vaidman, L. Causality constraints on nonlocal quantum measurements. *Phys. Rev. A* **49**, 4331–4338 (1994).
- Short, T. & Wehner, S. Entropy in general physical theories. *N. J. Phys.* **12**, 033023 (2010).
- Popescu, S. & Rohrlich, D. Quantum nonlocality as an axiom. *Found. Phys.* **24**, 379–385 (1994).
- Popescu, S. & Rohrlich, D. *The dilemma of Einstein, Podolsky and Rosen, 60 years later: International symposium in honour of Nathan Rosen* (eds Mann, A. & Revzen, M.) (Institute of Physics Pub., 1996).
- Popescu, S. & Rohrlich, D. *Proceedings of the Symposium of Causality and Locality in Modern Physics and Astronomy: Open Questions and Possible Solutions* (eds Hunter, G., Jeffers, S. & Vigier, J.-P.) 383 (Kluwer Academic Publishers, Dordrecht/Boston/London, 1997).
- Clauser, J., Horne, M., Shimony, A. & Holt, R. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880–884 (1969).
- Pawlowski, M. *et al.* Information causality as a physical principle. *Nature* **461**, 1101–1104 (2009).
- van Dam, W. *Nonlocality and Communication Complexity* PhD Thesis (University of Oxford, Department of Physics, 2000).
- Barnum, H., Beigi, S., Boixo, S., Elliot, M. & Wehner, S. Local quantum measurement and no-signaling imply quantum correlations. *Phys. Rev. Lett.* **104**, 140401 (2010).
- Oppenheim, J. & Wehner, S. The uncertainty principle determines the non-locality of quantum mechanics. *Science* **330**, 1072–1074 (2010).
- Chiribella, G., D'Ariano, G. M. & Perinotti, P. Informational derivation of quantum theory. *Phys. Rev. A* **84**, 012311 (2011).
- Masanes, L. & Müller, M. P. A derivation of quantum theory from physical requirements. *N. J. Phys.* **13**, 063001 (2011).

Acknowledgements

We thank Christian Gogolin, Paolo Perinotti, Marco Tomamichel and Markus Baden for insightful discussions and Matthew Pusey for interesting comments on a preliminary version of this work. This research was supported by the Ministry of Education and the National Research Foundation, Singapore.

Author contributions

Both authors contributed equally to the development of the main ideas. C.P. proved the technical claims and wrote most of the manuscript.

Additional information

Supplementary Information accompanies this paper at <http://www.nature.com/naturecommunications>

Competing financial interests: The authors declare no competing financial interests.

Reprints and permission information is available online at <http://npg.nature.com/reprintsandpermissions/>

How to cite this article: Pfister, C. & Wehner, S. An information-theoretic principle implies that any discrete physical theory is classical. *Nat. Commun.* 4:1851 doi: 10.1038/ncomms2821 (2013).