ARTICLE   OPEN

Quantum computation is the unique reversible circuit model for which bits are balls

Marius Krumm[1,2] and Markus P. Müller[1,3]

The computational efficiency of quantum mechanics can be characterized in terms of the qubit circuit model, which is defined by a few simple properties: each computational gate is a reversible transformation in a connected matrix group; single wires carry quantum bits, i.e. states of a three-dimensional Bloch ball; states on two or more wires are uniquely determined by local measurement statistics and their correlations. In this paper, we ask whether other types of computation are possible if we relax one of those characteristics (and keep all others), namely, if we allow wires to be described by d-dimensional Bloch balls, where d is different from three. Theories of this kind have previously been proposed as possible generalizations of quantum physics, and it has been conjectured that some of them allow for interesting multipartite reversible transformations that cannot be realized within quantum theory. However, here we show that all such potential beyond-quantum models of computation are trivial: if d is not three, then the set of reversible transformations consists entirely of single-bit gates, and not even classical computation is possible. In this sense, qubit quantum computation is an island in theoryspace.

INTRODUCTION

Since the discovery of quantum algorithms that outperform all known classical ones in certain tasks,[1] improving our understanding of the possibilities and limitations of quantum computation has become one of the central goals of quantum information theory. While it is notoriously difficult to prove unconditional separation of polynomial-time classical and quantum computation,[2] an approach that is often regarded more tractable is to analyze how certain modifications of quantum computing affect its computational power. For instance, one may consider restrictions on the set of allowed quantum resources, and ask under which condition the possibility of universal quantum computation is preserved despite the restriction. Notable results along these lines, among many others, include the Gottesman–Knill theorem,[3–5] insights on the necessity of contextuality as a resource for magic state distillation,[6] or bounds on the noise threshold of quantum computers.[7]

In a complementary and in some sense more radical approach, going back to Abrams and Lloyd,[8] one considers modifications of the quantum formalism itself and studies the impact of those modifications on the computational efficiency, resembling strategies of classical computer science such as the introduction of oracles.[9] For example, it has been shown that availability of closed timelike curves leads to implausible computational power,[10] that stronger-than-quantum nonlocality reduces the set of available transformations,[11–14] that tomographic locality forces computations to be contained in a class called AWPP,[15,16] and that in some theories (satisfying additional axioms) higher-order interference[17] does not lead to a speed-up in Grover's algorithm.[18] Further examples can be found, e.g., in refs. [19–22]

In this paper, we consider a specific modification of the quantum formalism that is arguably among the simplest and most conservative possibilities. This modification dates back to ideas by Jordan et al.,[23] and it has several independent motivations as we will explain further below. This generalization keeps all characteristic properties of quantum computation unchanged, but modifies a single aspect: namely, it allows the quantum bit to have any number of $d \geq 2$ degrees of freedom, instead of standard quantum theory's $d = 3$ (or the classical bit's $d = 1$). It has been conjectured[24] that the resulting theories allow for interesting "beyond quantum" reversible multipartite dynamics, which would make the corresponding models of computation highly relevant objects of study within the research program mentioned above. However, here we show that, quite on the contrary, these models are so constrained that they do not even allow for classical computation; hence, in Aaronson's terminology, the $d = 3$ case of the standard qubit circuit model can be seen as an "island in theoryspace".[25]

RESULTS

The results of this paper are organized as follows:

At first we motivate and explain the framework: In "Framework: Single gbits" section, we define single bits that generalize the qubit ("gbits"), and afterwards in "Framework: Gbit circuits" section, we give three postulates that allow us to reason about circuits that are constructed out of n of these gbits. We formulate the problem that is addressed in this work and describe how it relates to earlier results in the literature in "$d = 3$ equals quantum computation, and relation to earlier work" section. In "Main result" section, we state our main result: namely, while our principles uniquely determine quantum computation in the case that the

[1]Institute for Quantum Optics and Quantum Information, Austrian Academy of Sciences, Boltzmanngasse 3, 1090 Vienna, Austria; [2]Vienna Center for Quantum Science and Technology (VCQ), Faculty of Physics, University of Vienna, Boltzmanngasse 5, 1090 Vienna, Austria and [3]Perimeter Institute for Theoretical Physics, Waterloo, ON N2L 2Y5, Canada
Correspondence: Marius Krumm (marius.krumm@univie.ac.at) or Markus P. Müller (markusm23@univie.ac.at)
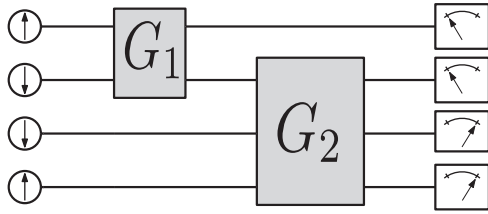
Fig. 1 The circuit model that we consider in this paper. We have an arbitrary finite number n of wires (here $n = 4$), and each wire carries a "gbit" which is a state in a d-dimensional Bloch ball state space. Initially, a product state is prepared (encoding, for example, the classical input to the algorithm), then a finite number of gates G_i is applied, each acting on an arbitrary number of gbits, and finally local measurements are performed. We assume that the G_i are elements of an (arbitrary unspecified) closed connected matrix group, and that the global state of n wires is uniquely determined by the statistics and correlations of single-wire measurements ("tomographic locality"). If $d = 3$, i.e., if the gbits are qubits, it has been shown in ref. [26] that these assumptions uniquely characterize unitary quantum computation as the only computationally nontrivial theory. Here we analyze the case $d \neq 3$, and prove that—despite conjectures to the opposite[24]—the corresponding models do not allow for any non-trivial computation at all. We do not assume that wires can be swapped, or that all transformations can be composed out of two-gbit transformations. See the main text for details

single gbits have $d = 3$ degree of freedom, any other value of d does not even allow for classical computation.

Framework: Single gbits

In both classical and quantum computation, we can restrict our attention to the circuit model (as in Fig. 1) where each of the wires (the single systems that enter and exit logical gates) corresponds to a two-level system. Quantum two-level systems (qubits) are different from classical ones (bits): they allow for a more complex behavior which encompasses phenomena like coherent superposition, interference, or uncertainty relations. Yet, both classical and quantum bits can be formalized in a unified way that we now describe (for both single and multiple bits, i.e., circuits, we follow the construction and notation from ref. [26]).

To any $d \in \mathbb{N}$, we associate a "generalized bit" (gbit) that has the d-dimensional Bloch ball, $B^d = \{\vec{a} \in \mathbb{R}^d \,|\, \|\vec{a}\| \leq 1\}$, as its state space. Every vector $\vec{a}$ in the Bloch ball B^d corresponds to a possible state of the generalized bit. Two-outcome measurements are described by vectors $\vec{b} \in \mathbb{R}^d$ with $|\vec{b}| = 1$, such that the probability of the first outcome if performed on state $\vec{a}$ is $(1 + \vec{a} \cdot \vec{b})/2$, and that of the second outcome is $(1 - \vec{a} \cdot \vec{b})/2$. In the following, it will be convenient to use the notation $v(\vec{a}) = (1, \vec{a})^\top \in \mathbb{R}^{d+1}$, such that these two probabilities become $\frac{1}{2} v(\vec{a}) \cdot v(\pm\vec{b})$. Reversible transformations of states are given by $\vec{a} \mapsto R\vec{a}$, where $R \in SO(d)$ is a rotation matrix. These transformations map states to states and can be inverted (by applying R^{-1}), hence we can interpret them as closed-system time evolutions or, equivalently, reversible gates on single generalized bits.

For $d = 3$, this formalism recovers the qubit of standard quantum theory:[5] as is well-known, every 2×2 density matrix ρ can be written in the form

$$\rho = (\mathbf{1} + \vec{a}_\rho \cdot \vec{\sigma})/2,$$

where $\vec{\sigma} = (\sigma_x, \sigma_y, \sigma_z)$ denotes the Pauli matrices. It is automatic in this representation that $\text{tr}\rho = 1$, and positivity $\rho \geq 0$ is equivalent to $|\vec{a}_\rho| \leq 1$. Hence the set of states of a quantum bit can be represented by the Bloch ball B^3. This representation has the important property that statistical mixtures correspond to convex

combinations: if a state ρ is prepared with probability p and another state ρ' is prepared with probability $1 - p$, then the total state $p\rho + (1-p)\rho'$ corresponds to the Bloch vector $\vec{a}_{p\rho+(1-p)\rho'} = p\vec{a}_\rho + (1-p)\vec{a}_{\rho'}$. This statistical interpretation of convex mixtures is also taken for balls of other dimensions $d \neq 3$, hence these Bloch balls can be regarded as state spaces of generalized probabilistic theories.[11]

In the $d = 3$ case, projective measurements are represented by unit vectors $\vec{b}$, $|\vec{b}| = 1$, with outcome probabilities $(1 \pm \vec{a} \cdot \vec{b})/2$ as described above. Unitary transformations U on states, acting as $\rho \mapsto U\rho U^\dagger$, are described in the Bloch ball picture by orthogonal maps R_U, $R_U^\top R_U = \mathbf{1}$, such that $\vec{a}_{U\rho U^\dagger} = R_U \vec{a}_\rho$. More general measurements (positive operator-valued measures) or transformations (completely positive maps) can also be described in the Bloch ball representation, but they are not needed in what follows and therefore omitted.

The simplest case of $d = 1$ corresponds to the classical bit: there are two possible configurations, $\vec{a} = +1$ and $\vec{a}' = -1$, and further states that represent classical uncertainty about the configuration. Namely, if we have $+1$ with probability p (and thus -1 with probability $1 - p$), this corresponds to the state $p\vec{a} + (1-p)\vec{a}'$ in the interior of the one-dimensional "Bloch ball".

There is one peculiarity in the $d = 1$ case: instead of $SO(1) = \{\mathbf{1}\}$, we should allow the group $O(1) = \{-\mathbf{1}, \mathbf{1}\}$ as Bloch ball transformations such that also the bit flip is allowed.

What is the significance of the d-dimensional Bloch balls if d is neither one nor three? These gbits have appeared in various places in quantum information theory and the foundations of quantum mechanics. Historically, they have first shown up as precisely those two-level state spaces that can be described as (formally real, irreducible) Jordan algebras,[23] a natural algebraic generalization of standard quantum theory. In fact, quantum theory with real amplitudes, i.e., over the field $\mathbb{R}$ instead of $\mathbb{C}$, has a $(d = 2)$-dimensional Bloch ball as its "quantum bit", and the bits of[27] quaternionic[28] and octonionic quantum theory correspond to B^d for $d = 5$ and $d = 9$, respectively. Furthermore, the fact that a two-level system should have a Euclidean ball state space can be derived from a variety of different sets of natural assumptions. In many reconstructions of quantum theory from physical or information-theoretic principles,[29–40] this fact is derived as a first step. For example, postulating that the group of reversible transformations acts transitively on the pure states implies that the pure states must all lie on the unit hypersphere of an invariant inner product. If some points on the sphere were not valid states, then there would exist additional measurements that would violate further natural postulates like Hardy's[29] "Subspaces" axiom. This argumentation or others along similar lines[29–40] lead to Euclidean balls as the most natural state spaces of a generalized bit.

A more geometrical motivation can be found by considering spin-$\frac{1}{2}$ particles (compare, e.g., to ref. [24]): under rotations $SO(3)$, they transform via $SU(2)$. The density matrix transforms under the adjoint representation, which means that the Bloch vectors transform via the same rotation as in physical space. Therefore, the Bloch vector $\vec{b}$ can be seen as defining an oriented axis in physical space. The model considered in this paper is a direct generalization of the Bloch ball and this interpretation to arbitrary spatial dimensions. Indeed, the possibility that space might have more than three dimensions has appeared in a large variety of physical theories.[41–46] It has also been argued that these generalized bits can be interpreted as "information quasiparticles" in some sense.[47] In summary, these gbits are among the simplest and most natural generalizations of the classical bit and the qubit of quantum mechanics.

Framework: Gbit circuits

To describe circuit computation, we need to define the state space, measurements, and transformations of several gbits. In standard quantum theory, where the gbits are qubits, there is a unique definition of these notions: the states of n qubits are exactly the $(2^n) \times (2^n)$ density matrices, the reversible transformations are the unitaries, and the measurements are described by collections of projection operators. Similar definitions apply to n classical bits. But if the gbits are Bloch balls of dimension $d \notin \{1, 3\}$, then it is apriori unclear what the composite state space should be.

Since we would like to be as general as possible, we will not make any attempt to fix the composite state space from the outset. Instead, we will work with a small set of principles that the composite n-gbit system is supposed to satisfy. While these principles will constrain the n-gbit state space, it is by no means obvious that they determine it uniquely. However, we will show below that they are indeed constraining enough to allow us to derive the full set of states and transformations.

An important principle is the no-signaling principle:[11] the outcome statistics of measurements on any group of gbits does not depend on any other operations (e.g., measurements) that are performed on the remaining gbits. This is a physically well-motivated constraint that lies at the heart of what we mean by "different wires" (i.e., subsystems) of the circuit in the first place.

This principle is satisfied by classical as well as quantum computation, and so is our second postulate of tomographic locality:[29,48] every state on n gbits is uniquely characterized by the statistics and correlations of the local gbit measurements. In other words, a global n-gbit state is nothing but a catalog of probabilities for the outcomes of all the single-gbit measurements and their correlations.

It is not only classical and quantum theory that satisfies the principle of tomographic locality, but also more general probabilistic theories like boxworld.[12] If this principle was violated, then a collection of gbits would in some counterintuitive sense be "more" than a composition of its building blocks. Even though this formulation makes tomographic locality sound very natural, there are simple examples of theories that violate it. One such example is given by quantum theory over the real numbers $\mathbb{R}$.[49,50] This is because observables of two single real qubits do not linearly generate all observables of two real qubits. In particular, if σ_y is the Pauli matrix with purely imaginary entries, then σ_y is not a real qubit observable, but $\sigma_y \otimes \sigma_y$ is a real two-qubit observable. Intuitively, it represents a novel "holistic" degree of freedom that cannot be constructed out of local degrees of freedom and their correlations.

Not only is the postulate of tomographic locality very intuitive, but it is also very powerful: it allows us to represent states of n gbits as tensors.[11] That is, even if we do not know what the set of n-gbit states is, we know that every such state can be written as an element of the linear space $\left(\mathbb{R}^{d+1}\right)^{\otimes n}$ (in the quantum case, where $d = 3$, this amounts to the 4^n-dimensional real linear space of Hermitian $(2^n) \times (2^n)$ matrices; for real bits, it is the 2^n-dimensional space that contains the probability vectors over 2^n configurations). In particular, an n-gbit product state with local Bloch vectors $\vec{a}_1, \ldots \vec{a}_n$ is represented by

$$v(\vec{a}_1, \ldots, \vec{a}_n) := (1, \vec{a}_1)^\top \otimes \ldots \otimes (1, \vec{a}_n)^\top,$$

and all other states ω are vectors on the same space (but not of this product form). Tomographic locality then amounts to the fact that all these states are uniquely determined by the numbers

$$2^{-n} v\left(\vec{b}_1, \ldots, \vec{b}_n\right)^\top \omega,$$

which are the outcome probabilities of local gbit measurements corresponding to the Bloch vectors $\vec{b}_1, \ldots, \vec{b}_n$ on the state ω. This mathematical property has many intuitively appealing consequences that are not otherwise guaranteed, e.g., the property that products of pure states are pure. It is also the reason why the mathematical literature has focused almost entirely on this notion of composite state space (cf., e.g., ref. [51]): it leads to notions of "tensor products" of ordered linear spaces that allow one to prove general statements that are otherwise unavailable. In the context of this paper, it would seem extremely difficult to make any meaningful statements whatsoever if not even the linear space on which the global states live could be fixed from the outset.

We need one further ingredient to arrive at a model of computation, namely a set of reversible transformations. In analogy to standard quantum computation (where these are the unitaries), we postulate that the transformations form a closed connected matrix group, and thus Lie group, $\mathcal{G}$: they form a group since they can be composed; they must be linear maps since if we prepare a state ω with probability p and ω' with probability $(1 - p)$, they must act on the components of the convex combination $p\omega + (1 - p)\omega'$ individually, to be consistent with the probabilistic interpretation.[11] Moreover, it is physically meaningful to model the group as closed since whenever we can approximate a transformation to arbitrary accuracy by gates, it makes sense to declare this transformation as in principle implementable.

This postulate is almost, but not quite, satisfied by classical computation, i.e., the $d = 1$ case. As Bennett has shown,[52] classical computation can be made fully reversible, at only marginal cost of space or time resources. There are finite universal gate sets (including, e.g., Toffoli gates) that generate the full group of permutations of the 2^n configurations of the n bits. These permutations therefore constitute the reversible transformations of the classical bits, and they form a closed matrix group of linear maps. This group, however, is discrete and not connected.

This discreteness is already reflected in the fact that the one-dimensional "Bloch ball" is discrete, i.e., has only a finite number (two) of pure states. Since the set of classical configurations (pure states) of n bits is discrete, the group of reversible transformations must also be discrete. In the case $d \geq 2$ to which we thus restrict our attention in the following, however, even single bits (Bloch balls) contain a continuous manifold of pure states. In order to allow every pure state to evolve into every other (which we would expect to be crucial for the exploitation of the full computational potential), it is therefore necessary that the reversible transformations form a continuous group $\mathcal{G}$—in more detail, that $\mathcal{G}$ is a matrix Lie group such that its connected component at the identity is non-trivial. It then makes sense to consider continuous time evolution that implements elements of this connected component (as it is the case in quantum theory), and to disregard the mathematical possibility of having additional disconnected components. This motivates the assumption that $\mathcal{G}$ is connected.

All gates in a circuit will be elements of $\mathcal{G}$. This group must in particular contain the local qubit rotations: for $R \in \mathrm{SO}(d)$, write $\hat{R}(1, \vec{a})^\top := (1, R\vec{a})^\top$, then the subgroup of local transformations is

$$\mathcal{G}_{\mathrm{loc}} := \left\{ \hat{R}_1 \otimes \hat{R}_2 \otimes \ldots \otimes \hat{R}_n \mid R_i \in \mathrm{SO}(d) \right\}.$$

Note that we have used tomographic locality in deriving this prescription: since a local transformation acts like a product of transformations on the product states, it must act like this on all other states too since they live on the vector space that is spanned by the product states. Tomographic locality hence enforces that we can represent any linear map $X : \left(\mathbb{R}^{(d+1)}\right)^{\otimes n} \to \left(\mathbb{R}^{(d+1)}\right)^{\otimes n}$ as a tensor with n upper and n lower indices; that is,

$$X^{a_1 a_2 \ldots a_n}_{\beta_1 \beta_2 \ldots \beta_n} := \left(\vec{e}_{\beta_1} \otimes \ldots \otimes \vec{e}_{\beta_n}\right)^\top X(\vec{e}_{a_1} \otimes \ldots \otimes \vec{e}_{a_n}),$$

where $0 \leq a_i, \beta_i \leq d$, and $\vec{e}_\gamma$ denotes the γ-th unit vector, e.g., $\vec{e}_0 = (1, 0, \ldots, 0)^\top$. This is in contrast to Bloch vectors $\vec{b} \in \mathbb{R}^d$, where we use the notation $\mathbb{R}^d \ni \vec{b} = \vec{e}_1 = (1, 0, \ldots, 0)^\top$.

We demand that $\mathcal{G}_{loc} \subseteq \mathcal{G}$, but do not make any further assumptions on $\mathcal{G}$. In particular, we do not assume that the n gbits have physically identical roles: our assumptions allow in principle composite state spaces of n gbits that are not symmetric with respect to permutations of the gbits. Hence we are also not assuming that gbits can be reversibly swapped, or that other natural choices of transformations such as extensions of classical reversible gates (like CNOT) can necessarily be implemented. Therefore, our framework does not rely on the same set of assumptions as the circuit framework of symmetric monoidal categories[53] that is often used in the quantum foundations context.

For any Bloch ball dimension d, there is a trivial computational model: namely the choice that $\mathcal{G} = \mathcal{G}_{loc}$. This describes a theory where the only possible reversible transformations are independent local transformations of the single gbits. Such a model does not even allow for classical gates like the CNOT; it only admits gates and computations that evolve the gbits independently from each other without ever correlating them, i.e., products of single-gbit gates. A state space that is compatible with this choice of global transformations is simply

$$\mathrm{conv}\left\{(1, \vec{a}_1)^\top \otimes \ldots \otimes (1, \vec{a}_n)^\top \,|\, \vec{a}_i \in B^d\right\},$$

i.e., all convex combinations of product states. This is a state space that does not contain entanglement.

$d = 3$ equals quantum computation, and relation to earlier work

For the case of the standard qubit, i.e., of $d = 3$, it has been proven in ref. [26] that there is only a single possible non-trivial ($\mathcal{G}_{loc} \subsetneq \mathcal{G}$) theory that satisfies the assumptions from above: namely, standard quantum theory over n qubits, with the $(2^n) \times (2^n)$ density matrices as the states, and the projective unitary group $\mathcal{G} = \mathrm{PU}(2^n)$ of transformations. That is, the postulates on composition of gbits from above, together with the structure of the single qubit, are sufficient to determine qubit quantum computation uniquely.

While this result is interesting in its own right, it is also the main motivation for the present work: if quantum computation is characterized by such a simple list of principles, then maybe one obtains other interesting models of computation by slightly tweaking one of the postulates. Since large parts of the mathematical structure are determined by the postulates on composition (no-signaling and tomographic locality), the most promising road towards modifying the setup and also keeping important mathematical tools seems to be to modify the structure of the single qubit—and technically as well as conceptually (as explained in "Framework: Single gbits" subsection), the most natural way to do this is by changing the dimension of the Bloch ball d.

In the special case of $n = 2$ gbits, the consequences of the above postulates have been explored in refs. [54,55]. There it has been proven that the only consistent choice of transformations for Bloch ball dimension $d \neq 3$ is given by the trivial choice $\mathcal{G} = \mathcal{G}_{loc}$. However, computation is typically taking place on a large number $n \gg 2$ of gbits, and the techniques of refs. [54,55] cannot readily be generalized to $n > 2$.

In fact, it has been suggested in ref. [24] that it is essential for Bloch ball dimensions $d \geq 4$ to allow for genuine m-partite interaction of the gbits, where $m \geq d - 1 \geq 3$. Without a conclusive proof or explicit construction of the state space, the authors conjectured that interesting multipartite reversible dynamics is possible for such systems. In contrast to quantum theory, this m-partite dynamics would not be decomposable into two-gbit interactions. While tomographic locality has not been assumed in ref. [24], it is an important first step to check their conjecture under this additional assumption. In fact, it has been argued in ref. [56] that in the context of spacetime physics (the Bloch balls are interpreted

in ref. [24] as carrying some sort of d-dimensional spin degrees of freedom), tomographic locality is to be expected due to arguments from group representation theory.

This gives us another, independent motivation to ask the main question of this paper: if $d \neq 3$ and n is any finite number of gbits, then what are the possible theories that satisfy the assumptions of "Framework: Gbit circuits" subsection?

Main result

The main result of this work is an answer to the question posed at the end of the previous section:

Theorem 1. *Consider a theory of n gbits, where single gbits are described by a $(d \geq 2)$-dimensional Bloch ball state space, subject to the single-gbit transformation group SO(d). As described above, let us assume no-signaling, tomographic locality, and that the global transformations form a closed connected matrix group $\mathcal{G}$. If $d \neq 3$, then necessarily $\mathcal{G} = \mathcal{G}_{loc}$, i.e., the only possible gates are (independent combinations of) single-gbit gates. No transformation can correlate gbits that are initially uncorrelated; hence not even classical computation is possible.*

Theorem 1 will be proved in "Methods" section.

DISCUSSION

Given a few simple properties that turn out to characterize qubit quantum computation, we have considered a natural modification: allowing the single bits to have more or less than the qubit's $d = 3$ degrees of freedom. We have analyzed the set of possible reversible transformations in the resulting theories, under the conjecture[24] (and in hopes) of discovering novel computational models that differ in interesting ways from quantum computation. Unfortunately, it turns out that the resulting models do not allow for any non-trivial reversible gates whatsoever. This reinforces earlier intuition[25] that quantum theory, or in this context quantum computation, is an "island in theoryspace".

While we have made an effort to be as careful and parsimonious in our assumptions as possible, it is still interesting to ask whether there are any remaining "loopholes" that could in principle leave some wiggle room for non-trivial beyond-quantum computation: can any of the assumptions of "Framework: Gbit circuits" subsection be dropped or weakened, while insisting that single bits are described by Bloch balls? We discuss several options in the Supplementary Material; in short, the most promising (but difficult) approaches would be to drop tomographic locality, and/or to drop reversibility or continuity of transformations. Both options present formidable mathematical challenges and are therefore deferred to future work.

The "rigidity" of quantum theory, i.e., the difficulty of modifying it in consistent ways, has been recognized in different contexts for a long time, see, e.g., Weinberg's proposal of a nonlinear modification of quantum mechanics,[57] and Gisin's subsequent discovery[58] that this modification allows for superluminal signaling. The research presented in this paper and in other work (like refs. [59,60]) makes this intuition more rigorous by specifying which combinations of principles already enforce the familiar behavior of quantum theory. These insights also illuminate our understanding of quantum computation, since they tell us which physical principles enforce its properties, and/or which other theoretical models of computation are plausibly conceivable.

Finally, it is interesting to speculate that the result of this paper is indirectly related to spacetime physics. After all, it is the fact that a qubit is represented as a 3-ball B^3, with SO(3) as its transformation group, which allows for spin-1/2 particles that couple to rotations in three-dimensional space. Given the popularity of approaches in which spacetime emerges in some

way from an underlying quantum theory,[61–63] this observation can perhaps be regarded as more than a coincidence. In fact, it has been argued more rigorously that the structures of quantum theory and spacetime mutually constrain each other.[24,56,64–66] This suggests a slogan that also fits some other ideas from quantum information:[67] the limits of computation are the limits of our world.

METHODS

We will now prove this result for the case $d \geq 4$. The proof in the $d = 2$ case uses similar techniques, but differs in several details for group-theoretic reasons. It will hence be deferred to the Supplementary Material.

Generator normal form for all dimensions $d \geq 2$

As a first step, we will consider the generators of global transformations and show that there exists at least one that is of a certain normal form. This part of the proof is valid for all dimensions $d \geq 2$. A large part of this first step follows the construction in ref. [26], and extends it to arbitrary dimensions.

Let $G \in \mathcal{G}$ be a transformation of the composite system. Suppose we prepare n gbits initially in states with Bloch vectors $\vec{a}_1, \ldots, \vec{a}_n$, evolve the resulting product state via G, and perform a final local n-gbit measurement with Bloch vectors $\vec{b}_1, \ldots, \vec{b}_n$. The probability that the all the n outcomes on the n gbits are "yes" is

$$2^{-n} v\left(\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_n\right)^\top G v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \in [0, 1].$$

Let us consider a group element $G = e^{\epsilon X}$ with $X \in \mathfrak{g}$ (the corresponding Lie algebra) and $\epsilon \in \mathbb{R}$ and expand:

$$v\left(\vec{b}_1, \ldots, \vec{b}_n\right)^\top \left(\mathbf{1} + \epsilon X + \frac{\epsilon^2}{2} X^2 + \mathcal{O}(\epsilon^3)\right) v(\vec{a}_1, \ldots, \vec{a}_n) \in [0, 2^n].$$

From now on we restrict ourselves to unit length Bloch vectors, i.e., $|\vec{a}_i| = |\vec{b}_j| = 1$ for all i, j. We obtain

$$\mathcal{C}[\vec{a}_1] := v\left(-\vec{a}_1, \vec{b}_2, \ldots, \vec{b}_n\right)^\top X v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = 0$$

since the zeroth order is zero which is a local minimum as a function of ϵ (see Fig. 2 for further explanation). Thus the second order contribution has to be non-negative:

$$v\left(-\vec{a}_1, \vec{b}_2, \ldots, \vec{b}_n\right)^\top X^2 v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \geq 0,$$

or more generally with the roles of qubits 1 and k exchanged,

$$v\left(\vec{b}_1, \ldots, \vec{b}_{k-1}, -\vec{a}_k, \vec{b}_{k+1}, \ldots \vec{b}_n\right)^\top X^2 v(\vec{a}_1, \ldots, \vec{a}_n) \geq 0. \quad (1)$$
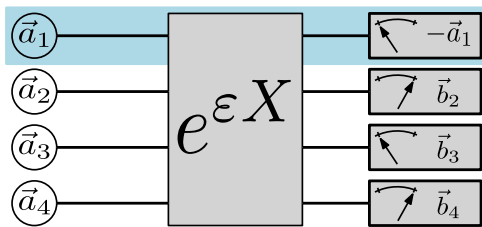

Fig. 2 We are using configurations like this one to derive constraints on the generators $X \in \mathfrak{g}$. In the special case $\epsilon = 0$, the transformation $\exp(\epsilon X)$ reduces to the identity. Hence, if we prepare the first wire in the (pure) state with Bloch vector $\vec{a}_1$, and perform a final measurement of that wire with Bloch vector $-\vec{a}_1$, the corresponding outcome will have probability zero, regardless of which local measurements we choose for the other wires. But probability zero is a local minimum, which implies that the derivative of this probability with respect to ϵ must be zero (yielding $C[\vec{a}_1] = 0$), and the second derivative must be non-negative (yielding constraint (1) in the case $k = 1$)

Other first and second order constraints are

$$v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)^\top X v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) = 0, \quad (2)$$

$$v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n)^\top X^2 v(\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_n) \leq 0 \quad (3)$$

for analogous reasons as above (since $\vec{b}_j = \vec{a}_j$ for all j yields probability one for $\epsilon = 0$, which is the global and thus a local maximum). For fixed Bloch vectors $\vec{a}_2, \ldots, \vec{a}_n, \vec{b}_2, \ldots, \vec{b}_n$, define W^α_β as

$$\left[\vec{e}_\beta \otimes \begin{pmatrix} 1 \\ \vec{b}_2 \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 \\ \vec{b}_n \end{pmatrix}\right]^\top X \left[\vec{e}_\alpha \otimes \begin{pmatrix} 1 \\ \vec{a}_2 \end{pmatrix} \otimes \ldots \otimes \begin{pmatrix} 1 \\ \vec{a}_n \end{pmatrix}\right]. \quad (4)$$

The equation $\mathcal{C}[\vec{e}_i] = 0$ implies $W^0_0 + W^i_0 - W^0_i - W^i_i = 0$, and $\mathcal{C}[-\vec{e}_i] = 0$ implies $W^0_0 - W^i_0 + W^0_i - W^i_i = 0$. Thus, $W^i_i = W^0_0$ and $W^0_i = W^i_0$ for all $i \geq 1$. Since the vectors $\begin{pmatrix} 1 \\ \vec{a} \end{pmatrix}$ linearly span all of $\mathbb{R}^{d+1}$, we get

$$X^{i\,a_2\,\ldots\,a_n}_{i\,\beta_2\,\ldots\,\beta_n} = X^{0\,a_2\,\ldots\,a_n}_{0\,\beta_2\,\ldots\,\beta_n}, \quad (5)$$

$$X^{i\,a_2\,\ldots\,a_n}_{0\,\beta_2\,\ldots\,\beta_n} = X^{0\,a_2\,\ldots\,a_n}_{i\,\beta_2\,\ldots\,\beta_n} \quad (6)$$

for all $i \geq 1$ and all $a_2, \ldots, a_n, \beta_2, \ldots, \beta_n \geq 0$. Similarly, $\mathcal{C}\left[\frac{1}{\sqrt{2}}(\vec{e}_i + \vec{e}_j)\right] = 0$ for $i \neq j$, $i, j \geq 1$ yields

$$W^0_0 + \frac{1}{\sqrt{2}} W^i_0 + \frac{1}{\sqrt{2}} W^j_0 - \frac{1}{\sqrt{2}} W^0_i - \frac{1}{2} W^i_i - \frac{1}{2} W^j_i - \frac{1}{\sqrt{2}} W^0_j - \frac{1}{2} W^i_j - \frac{1}{2} W^j_j = 0.$$

Using the results on W^i_i and W^0_i further above, this reduces to $-\frac{1}{2} W^j_i - \frac{1}{2} W^i_j = 0$, and thus

$$X^{i\,a_2\,\ldots\,a_n}_{j\,\beta_2\,\ldots\,\beta_n} = -X^{j\,a_2\,\ldots\,a_n}_{i\,\beta_2\,\ldots\,\beta_n} \quad (7)$$

for all $i, j \geq 1$ and $a_2, \ldots, a_n, \beta_2, \ldots, \beta_n \geq 0$. While we have derived (5), (6), and (7) for the first gbit, analogous equations hold for all other gbits with labels $2, \ldots, n$.

Let us denote by $\mathcal{A}$ the antisymmetric $(d + 1) \times (d + 1)$-matrices of the form

$$\mathcal{A} := \left\{ A_{\bar{A}} = \left(\begin{array}{c|c} 0 & \vec{0}^\top \\ \hline \vec{0} & \bar{A} \end{array}\right) \;\middle|\; \bar{A}^\top = -\bar{A} \right\},$$

and by $\mathcal{B}$ the symmetric $(d + 1) \times (d + 1)$-matrices of the form

$$\mathcal{B} := \left\{ B_{\vec{b}} = \left(\begin{array}{c|c} 0 & \vec{b}^\top \\ \hline \vec{b} & 0 \end{array}\right) \;\middle|\; \vec{b} \in \mathbb{R}^d \right\}.$$

Furthermore, let $\mathcal{I} := \mathbb{R} \cdot \mathbf{1}$, i.e., all multiples of the $(d + 1) \times (d + 1)$ identity matrix. The sets $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{I}$ are real linear matrix subspaces. Note that these three spaces are pairwise orthogonal with respect to the Hilbert–Schmidt inner product $\langle X, Y \rangle := \mathrm{tr}(X^\top Y)$. The matrix W defined in (4) must then be an element of $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$ due to the identities for its components that we have derived above. More generally, since the same identities hold for every index $i \in \{1, \ldots, n\}$ for the tensor X, we obtain $X \in (\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n}$. Since $X \in \mathfrak{g}$ was arbitrary, this tells us that

$$\mathfrak{g} \subset (\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n}.$$

The Lie algebra of the local transformations is

$$\mathfrak{g}_{\mathrm{loc}} = \mathcal{A} \otimes \mathbf{1} \otimes \ldots \otimes \mathbf{1} + \mathbf{1} \otimes \mathcal{A} \otimes \mathbf{1} \otimes \ldots \otimes \mathbf{1} + \ldots + \mathbf{1} \otimes \mathbf{1} \otimes \ldots \otimes \mathbf{1} \otimes \mathcal{A},$$

writing "+" instead of "$\oplus$" for readability. We can write the space $(\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n}$ in a somewhat different form. To this end, consider strings of symbols $x \in \{A, B, I\}^n$, for example, $x = ABAI$ (if $n = 4$), and denote the corresponding tensor product matrix spaces by S_x; for this example, $S_x = \mathcal{A} \otimes \mathcal{B} \otimes \mathcal{A} \otimes \mathcal{I}$. Then $S_x \perp S_y$ for $x \neq y$ (with respect to the Hilbert–Schmidt inner product), and

$$(\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I})^{\otimes n} = \bigoplus_{x \in \{A,B,I\}^n} S_x.$$

Now let $X \in \mathfrak{g} \setminus \mathfrak{g}_{\mathrm{loc}}$ be an arbitrary generator which is not in the local Lie algebra (here we explicitly make the assumption that such an X exists). Since $X \neq 0$, there must exist x such that $\Phi_x(X) \neq 0$ for the orthogonal

projection Φ_x into S_x, and since $X \notin \mathfrak{g}_{\text{loc}'}$, at least one of those x must satisfy

$$x \notin \{AI \ldots I, IAI \ldots I, \ldots, I \ldots IA\}.$$

Reordering the gbits, we may assume that $x = A^{n_A} B^{n_B} I^{n_I}$, where $n_A + n_B + n_I = n$ and one of the following three cases applies:

(i) $n_A = 0$,
(ii) $n_A = 1$ and $n_B \geq 1$,
(iii) $n_A \geq 2$,

Since S_x has an orthonormal basis of matrices of the form $A_{\bar{A}_1} \otimes \ldots \otimes A_{\bar{A}_{n_A}} \otimes B_1 \otimes \ldots \otimes B_{n_B} \otimes \mathbf{1}^{\otimes n_I}$, where all $A_{\bar{A}_i} \in \mathcal{A}$ and $B_i \in \mathcal{B}$, there must exist some matrix $\tilde{M}_x$ of that form (i.e., $\tilde{M}_x \in S_x$) such that $\langle X, \tilde{M}_x \rangle \neq 0$. By moving constant scalar factors into the A-terms, we may assume that there are unit vectors $\vec{b}_i$ such that $B_i = B_{\vec{b}_i}$ for $i = 1, \ldots, n_B$. But since $\hat{R} B_{\vec{b}} \hat{R}^\top = B_{R\vec{b}}$ for all $R \in SO(d)$, there are orthogonal matrices $\hat{R}_i$ such that $R_i \vec{b}_i = \vec{e}_1 = (1, 0, \ldots, 0)^\top$ for all i, and the local transformation $T := \mathbf{1}^{\otimes n_A} \otimes \hat{R}_1 \otimes \ldots \otimes \hat{R}_{n_B} \otimes \mathbf{1}^{\otimes n_I}$ satisfies

$$M'_x := T \tilde{M}_x T^{-1} = T \tilde{M}_x T^\top = A_{\bar{A}_1} \otimes \ldots \otimes A_{\bar{A}_{n_A}} \otimes B^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I},$$

where $B := B_{\vec{e}_1}$. Set $X' := TXT^{-1}$, then since $T \in \mathcal{G}_{\text{loc}} \subset \mathcal{G}$, and since the adjoint action of $\mathcal{G}_{\text{loc}}$ preserves $\mathfrak{g}_{\text{loc}}$, we have $X' \in \mathfrak{g} \backslash \mathfrak{g}_{\text{loc}}$, and $\langle X', M'_x \rangle = \text{tr}(TXT^{-1}T\tilde{M}_xT^{-1}) = \langle X, \tilde{M}_x \rangle \neq 0$. Similar argumentation allows us to bring the $A_{\bar{A}_i}$ into a standard form. Since the $d \times d$-matrices $\bar{A}_i$ are antisymmetric, one can infer from the results in refs. [17,27,28] that there are orthogonal transformations $R_i \in SO(d)$ such that

$$R_i \bar{A}_i R_i^\top = \begin{pmatrix} 0 & \lambda_1^{(i)} & & & & & \\ -\lambda_1^{(i)} & 0 & & & & & \\ & & 0 & \lambda_2^{(i)} & & & \\ & & -\lambda_2^{(i)} & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & \lambda_{d/2}^{(i)} \\ & & & & & -\lambda_{d/2}^{(i)} & 0 \end{pmatrix} \quad (d \text{ even}),$$

$$\begin{pmatrix} 0 & & & & & & \\ & 0 & \lambda_1^{(i)} & & & & \\ & -\lambda_1^{(i)} & 0 & & & & \\ & & & 0 & \lambda_2^{(i)} & & \\ & & & -\lambda_2^{(i)} & 0 & & \\ & & & & & \ddots & \\ & & & & & 0 & \lambda_{\frac{d-1}{2}}^{(i)} \\ & & & & & -\lambda_{\frac{d-1}{2}}^{(i)} & 0 \end{pmatrix} \quad (d \text{ odd}).$$

To save space, we will use the following notation in the remainder of the paper, where $\sigma = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$:

$$R_i \bar{A}_i R_i^\top = \begin{cases} \lambda_1^{(i)} \sigma \oplus \lambda_2^{(i)} \sigma \oplus \ldots \oplus \lambda_{d/2}^{(i)} \sigma & (d \text{ even}), \\ 0_{1 \times 1} \oplus \lambda_1^{(i)} \sigma \oplus \lambda_2^{(i)} \sigma \oplus \ldots \oplus \lambda_{\frac{d-1}{2}}^{(i)} \sigma & (d \text{ odd}). \end{cases}$$

Now consider the corresponding $(d+1) \times (d+1)$-matrices $A_{R_i \bar{A}_i R_i^\top}$, for which we will introduce the following notation. By A_j, denote the matrix for which only the jth block is non-zero, with $\lambda_j = 1$. That is, for even d, we have the $(d+1) \times (d+1)$-matrices

$$A_1 = 0_{1 \times 1} \oplus \sigma \oplus 0_{2 \times 2} \oplus \ldots \oplus 0_{2 \times 2},$$
$$A_2 = 0_{1 \times 1} \oplus 0_{2 \times 2} \oplus \sigma \oplus 0_{2 \times 2} \oplus \ldots \oplus 0_{2 \times 2},$$
$$\vdots$$
$$A_{d/2} = 0_{1 \times 1} \oplus 0_{2 \times 2} \oplus \ldots \oplus 0_{2 \times 2} \oplus \sigma,$$

and for odd d, we have an extra initial zero, namely

$$A_1 = 0_{2 \times 2} \oplus \sigma \oplus 0_{2 \times 2} \oplus \ldots \oplus 0_{2 \times 2},$$
$$A_2 = 0_{2 \times 2} \oplus 0_{2 \times 2} \oplus \sigma \oplus 0_{2 \times 2} \oplus \ldots \oplus 0_{2 \times 2},$$
$$\vdots$$
$$A_{(d-1)/2} = 0_{2 \times 2} \oplus 0_{2 \times 2} \oplus \ldots \oplus 0_{2 \times 2} \oplus \sigma.$$

The local transformation $\tilde{T} := \hat{R}_1 \otimes \ldots \otimes \hat{R}_{n_A} \otimes \mathbf{1}^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I}$ satisfies

$$M_x := \tilde{T} M'_x \tilde{T}^{-1} = \tilde{T} M'_x \tilde{T}^\top$$
$$= \left(\sum_j \lambda_j^{(1)} A_j \right) \otimes \ldots \otimes \left(\sum_j \lambda_j^{(n_A)} A_j \right) \otimes B^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I}, \quad (8)$$

where the $\lambda_j^{(i)}$ are real numbers. Set $X'' := \tilde{T} X' \tilde{T}^{-1}$, then since $\tilde{T} \in \mathcal{G}_{\text{loc}} \subset \mathcal{G}$, we have $X'' \in \mathfrak{g} \backslash \mathfrak{g}_{\text{loc}}$, and $\langle X'', M_x \rangle = \text{tr}(\tilde{T} X' \tilde{T}^{-1} \tilde{T} M'_x \tilde{T}^{-1}) = \langle X', M'_x \rangle \neq 0$.

In summary, we have shown that if there exist any nonlocal generators at all, then there is one (denoted X'') that has non-zero overlap with a matrix $M_x \in S_x$ of the simple form (8).

Next we will show that this implies that $\mathfrak{g} = \mathfrak{g}_{\text{loc}}$ for all Bloch ball dimensions $d \geq 4$.

Proof of Theorem 1 for $d \geq 4$

We now use Schur's Lemma to construct orthogonal projectors (with respect to the Hilbert–Schmidt inner product) onto the subspaces of $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$. First, define

$$\Phi_I[M] := \int_{SO(d)} \hat{R} M \hat{R}^{-1} dR \quad (M \in \mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}),$$

then $\Phi_I[M] = 0$ for all $M \in \mathcal{A} \oplus \mathcal{B}$ and $\Phi_I[M] = M$ for all $M \in \mathcal{I}$. Since these subspaces are orthogonal with respect to the Hilbert–Schmidt inner product, Φ_I is the orthogonal projector onto the subspace $\mathcal{I}$ of $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$ (we are not interested in its action on matrices that are not in the space $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$).

For $j = 1, \ldots, d$, consider the stabilizer subgroup

$$\mathcal{G}_j := \{R \in SO(d) | R\vec{e}_j = \vec{e}_j\},$$

where $\vec{e}_j$ denotes the jth standard unit vector in $\mathbb{R}^d$. Every $\mathcal{G}_j$ is isomorphic to $SO(d-1)$ whose fundamental representation is irreducible (note that this is not true for $d = 3$; this causes the crucial difference to ref. [26]). Set

$$\Phi_{\vec{e}_j}[M] := \int_{\mathcal{G}_j} \hat{R} M \hat{R}^{-1} dR \quad (M \in \mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}),$$

then $\Phi_{\vec{e}_1}[M] = \int_{SO(d-1)} \begin{pmatrix} \mathbf{1}_2 & \\ & S \end{pmatrix} M \begin{pmatrix} \mathbf{1}_2 & \\ & S^{-1} \end{pmatrix} dS$, and, similarly as above, Schur's Lemma implies that $\Phi_{\vec{e}_1}$ is the orthogonal projector onto $\text{span}(B) \oplus \mathcal{I}$. Hence $\Phi_B := \Phi_{\vec{e}_1} - \Phi_I$ is the orthogonal projector onto $\text{span}(B)$.

Finally, we will construct the orthogonal projector onto $\mathcal{A}_{\text{blocks}} := \text{span}\{A_1, \ldots, A_z\}$, where $z = d/2$ if d is even and $z = (d-1)/2$ if d is odd. To this end, define the $SO(2)$-matrix $R(\theta) := \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix}$, and set

$$\hat{R}(\theta_1, \theta_2, \ldots, \theta_z) := \begin{pmatrix} \mathbf{1}_y & & & \\ & R(\theta_1) & & \\ & & \ddots & \\ & & & R(\theta_z) \end{pmatrix},$$

where $y = 1$ if d is even and $y = 2$ if d is odd. Furthermore, define $\Phi'[M]$ as

$$\int_0^{2\pi} \frac{d\theta_1}{2\pi} \int_0^{2\pi} \frac{d\theta_2}{2\pi} \ldots \int_0^{2\pi} \frac{d\theta_z}{2\pi} \hat{R}(\theta_1, \ldots, \theta_z) M \hat{R}(\theta_1, \ldots, \theta_z)^{-1}.$$

Using the identities

$$\int_0^{2\pi} R(\theta) \frac{d\theta}{2\pi} = 0,$$
$$\int_0^{2\pi} R(\theta) \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} R(-\theta) \frac{d\theta}{2\pi} = \frac{1}{2} \begin{pmatrix} m_{11} + m_{22} & m_{12} - m_{21} \\ -m_{12} + m_{21} & m_{11} + m_{22} \end{pmatrix}$$
$$=: \Psi \left[\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \right].$$

we can evaluate the action of Φ' as follows. First, any given $(d+1) \times (d+$

1)-matrix M can be written in the block matrix form

$$M = \begin{pmatrix} M_{0,0} & & M_{0,z} \\ \vdots & \ddots & \vdots \\ M_{z,0} & & M_{z,z} \end{pmatrix}$$

where $M_{0,0}$ is a $y \times y$-matrix, all $M_{i,j}$ for $i, j \geq 1$ are 2×2-matrices, and the other matrices are $y \times 2$ and $2 \times y$-matrices. Then, the action of Φ' becomes

$$\Phi'[M] = \begin{pmatrix} M_{0,0} & 0 & & 0 \\ 0 & \Psi[M_{1,1}] & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & & 0 & \Psi[M_{z,z}] \end{pmatrix}.$$

Hence Φ' is an orthogonal projection that acts as the identity on $\mathcal{I}$ (i.e., Φ' $(\mathbf{1}) = \mathbf{1}$), and it projects $\mathcal{A}$ into its subspace $\mathcal{A}_{blocks}$. Furthermore, if d is even, then Φ' annihilates $\mathcal{B}$, and if d is odd, then Φ' projects $\mathcal{B}$ into its subspace span(B). Thus, for d even, the orthogonal projector onto $\mathcal{A}_{blocks}$ is $\Phi_A := \Phi' - \Phi_I$, and for d odd, it is $\Phi_A := \Phi' - \Phi_I - \Phi_B$. Note that all these statements are only claimed to hold for the case that the maps are applied to operators in $\mathcal{A} \oplus \mathcal{B} \oplus \mathcal{I}$.

The projectors Φ_I, Φ_B and Φ_A map the Lie algebra $\mathfrak{g}$ into itself, if we apply different products of those projectors to the n sites. For example, consider the special case $n = 1$. Then $Z \in \mathfrak{g}$ implies $\Phi_I[Z] \in \mathfrak{g}$ since $\mathfrak{g}$ is closed with respect to conjugations by elements of $\mathcal{G}$ and integrals. Similarly, $\Phi_{\vec{e}_1}[Z] \in \mathfrak{g}$, and since $\mathfrak{g}$ is a linear space, we also have $\Phi_B[Z] = \Phi_{\vec{e}_1}[Z] - \Phi_I[Z] \in \mathfrak{g}$, and similarly for the projector Φ_A. If $n \geq 2$, then we can successively apply the projectors to one of the sites, using the fact that tensoring local rotations with identities gives local transformations in $\mathcal{G}_{loc}$. Thus, if we define

$$\Phi := \Phi_A^{\otimes n_A} \otimes \Phi_B^{\otimes n_B} \otimes \Phi_I^{\otimes n_I},$$

then $Y := \Phi[X'']$ is another valid generator, $Y \in \mathfrak{g}$. Furthermore, $\Phi[M_x] = M_x$, hence

$$0 \neq \langle X'', M_x \rangle = \langle X'', \Phi[M_x] \rangle = \langle \Phi[X''], M_x \rangle = \langle Y, M_x \rangle \qquad (9)$$

and thus $Y \neq 0$ (we have used that Φ is an orthogonal projection and thus in particular self-adjoint with respect to the Hilbert–Schmidt inner product). In particular, $Y \in \text{Im}(\Phi) = \mathcal{A}_{blocks}^{\otimes n_A} \otimes \text{span}(B)^{\otimes n_B} \otimes \mathcal{I}^{\otimes n_I}$. Consequently, there are real numbers $\lambda_{j_1, \dots, j_{n_A}}$ such that

$$Y = \sum_{j_1, \dots, j_{n_A} = 1}^{z} \lambda_{j_1, \dots, j_{n_A}} A_{j_1} \otimes \dots \otimes A_{j_{n_A}} \otimes B^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I}.$$

Now we apply the identities $A_j A_k = -\delta_{jk} P_j$ and $B^2 = P_B$, where

$$P_B = \mathbf{1}_{2 \times 2} \oplus 0_{(d-1) \times (d-1)},$$
$$P_1 = 0_{y \times y} \oplus \mathbf{1}_{2 \times 2} \oplus 0_{2(z-1) \times 2(z-1)},$$
$$P_2 = 0_{y \times y} \oplus 0_{2 \times 2} \oplus \mathbf{1}_{2 \times 2} \oplus 0_{2(z-2) \times 2(z-2)}$$

and so on, up to P_z. This gives us

$$Y^2 = (-1)^{n_A} \sum_{j_1, \dots, j_{n_A}} \lambda_{j_1, \dots, j_{n_A}}^2 P_{j_1} \otimes \dots \otimes P_{j_{n_A}} \otimes P_B^{\otimes n_B} \otimes \mathbf{1}^{\otimes n_I}. \qquad (10)$$

Suppose that n_A is even so that $(-1)^{n_A} = 1$. We will now show that constraint (3) gets violated. To this end, fix some $j_1^0, \dots, j_{n_A}^0$ such that $\lambda_{j_1^0, \dots, j_{n_A}^0} \neq 0$. For $i = 1, \dots, n_A$, choose some unit vector $\vec{a}_i \in \mathbb{R}^d$ such that $\begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}^\top P_{j_i^0} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} > 0$; for all other j_i, we automatically get $\begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}^\top P_{j_i} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} \geq 0$. For $i = n_A + 1, \dots, n_A + n_B$, set $\vec{a}_i := \vec{e}_1$, then $\begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}^\top P_B \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} = 2$. Finally, for $i \geq n_A + n_B + 1$, choose $\vec{a}_i$ arbitrarily such that $\begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}^\top \mathbf{1} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} = 2$. Altogether, we obtain

$$v(\vec{a}_1, \dots, \vec{a}_n)^\top Y^2 v(\vec{a}_1, \dots, \vec{a}_n) > 0$$

which violates constraint (3). Thus n_A must be odd, and $(-1)^{n_A} = -1$.

Recall constraint (1) in the special case $k = 2$:

$$v\left(\vec{b}_1, -\vec{a}_2, \vec{b}_3, \dots, \vec{b}_n\right)^\top Y^2 v(\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n) \geq 0 \qquad (11)$$

for all unit vectors $\vec{a}_i, \vec{b}_j \in \mathbb{R}^d$. For all $i \in [n_A + n_B + 1, n] \backslash \{2\}$, choose $\vec{a}_i, \vec{b}_i$ such that $\begin{pmatrix} 1 \\ \vec{b}_i \end{pmatrix}^\top \mathbf{1} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} > 0$ (simply avoid the choice $\vec{a}_i = -\vec{b}_i$). Similarly, for all $i \in [n_A + 1, n_A + n_B] \backslash \{2\}$, choose $\vec{a}_i, \vec{b}_i$ such that $\begin{pmatrix} 1 \\ \vec{b}_i \end{pmatrix}^\top P_B \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix} > 0$. We will now distinguish two cases for n_A.

First, consider the case $n_A = 1$. Since our original generator X was chosen nonlocal, it follows that $n_B \geq 1$, as explained in "Generator normal form for all dimensions $d \geq 2$" subsection. Thus, the second tensor factor in (10) must be P_B. We will now choose $\vec{a}_2 = \vec{e}_2$ which implies that $\begin{pmatrix} 1 \\ -\vec{a}_2 \end{pmatrix}^\top P_B \begin{pmatrix} 1 \\ \vec{a}_2 \end{pmatrix} = 1$. But then we may still choose $\vec{b}_1, \vec{a}_1$ arbitrarily, and by choosing these two unit vectors suitably from the subspace $\text{Im}(P_{j_1})$, we may generate an arbitrary sign for $\begin{pmatrix} 1 \\ \vec{b}_1 \end{pmatrix}^\top P_{j_1} \begin{pmatrix} 1 \\ \vec{a}_1 \end{pmatrix}$. Thus, we can break constraint (11) by a suitable choice of these two unit vectors, which yields a contradiction.

Second, suppose that $n_A \geq 3$ (we already know that n_A must be odd). Then we can choose $\vec{a}_2$ such that $\begin{pmatrix} 1 \\ -\vec{a}_2 \end{pmatrix}^\top P_{j_2} \begin{pmatrix} 1 \\ \vec{a}_2 \end{pmatrix} = -1$. We have even more freedom than in the previous case: for all $i \in [1, n_A] \backslash \{2\}$, we can choose $\vec{b}_i, \vec{a}_i$ from the subspace $\text{Im}(P_{j_i})$ such that we get an arbitrary sign for every $\begin{pmatrix} 1 \\ \vec{b}_i \end{pmatrix}^\top P_{j_i} \begin{pmatrix} 1 \\ \vec{a}_i \end{pmatrix}$. This also leads to a violation of constraint (11), and we obtain a contradiction as well.

This means that our initial assumption must have been wrong—namely, that there exists a generator in $\mathfrak{g} \backslash \mathfrak{g}_{loc}$. We conclude that instead this set must be empty, hence $\mathfrak{g} = \mathfrak{g}_{loc}$. But since $\mathcal{G}$ is compact and connected, it follows from ref. [68] [Theorem VII.2.2 (v)] that $\mathcal{G}$ cannot be larger than $\mathcal{G}_{loc}$. This proves our main result, Theorem 1, for Bloch ball dimensions $d \geq 4$. The proof for $d = 2$ is given in the Supplementary Material.

AUTHOR CONTRIBUTIONS

M.P.M. contributed the original idea for the project. M.K. proved most of the technical results under the supervision of M.P.M. M.P.M. wrote the first draft of the paper, and both authors contributed to the final version.

ADDITIONAL INFORMATION

REFERENCES

1. Shor, P. W. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM J. Sci. Statist. Comput.* **26**, 1484 (1997).
2. Aaronson, S. *Quantum computing since democritus* (Cambridge University Press, New York, 2013).
3. Gottesman, D. Talk at the international conference on group theoretic methods in physics. Preprint at arXiv:quant-ph/9807006 (1998).

4. Aaronson, S. & Gottesman, D. Improved simulation of stabilizer circuits. *Phys. Rev. A* **70**, 052328 (2004).

5. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

6. Howard, M., Wallman, J., Veitch, V. & Emerson, J. Contextuality supplies the 'magic' for quantum computation. *Nature* **510**, 351–355 (2014).

7. Gottesman, D. An introduction to quantum error correction and fault-tolerant quantum computation. Preprint at arXiv:0904.2557 (2009).

8. Abrams, D. S. & Lloyd, S. Nonlinear quantum mechanics implies polynomial-time solution for NP-complete and sharp-P problems. *Phys. Rev. Lett.* **81**, 3992 (1998).

9. Sipser, M. *Introduction to the theory of computation* (PWS Publishing, Boston, MA, 1997).

10. Aaronson, S. & Watrous, J. Closed timelike curves make quantum and classical computing equivalent. *Proc. R. Soc. A* **465**, 631–647 (2009).

11. Barrett, J. Information processing in generalized probabilistic theories. *Phys. Rev. A* **75**, 032304 (2007).

12. Gross, D., Müller, M., Colbeck, R. & Dahlsten, O. C. O. All reversible dynamics in maximally nonlocal theories are trivial. *Phys. Rev. Lett.* **104**, 080402 (2010).

13. Al-Safi, S. W. & Short, A. J. Reversible dynamics in strongly non-local Boxworld systems. *J. Phys. A* **47**, 325303 (2014).

14. Richens, J. G., Selby, J. H. & Al-Safi, S. W. Entanglement is necessary for emergent classicality in all physical theories. *Phys. Rev. Lett.* **119**, 080503 (2017).

15. Lee, C. M. & Barrett, J. Computation in generalised probabilistic theories. *New J. Phys.* **17**, 083001 (2015).

16. Barrett, J., de Beaudrap, N., Hoban, M. J. & Lee, C. M. The computational landscape of general physical theories. Preprint at arXiv:1702.08483 (2017).

17. Lee, C. M. & Selby, J. H. Deriving Grover's lower bound from simple physical principles. *New J. Phys.* **18**, 093047 (2016).

18. Barnum, H., Lee, C. M., Scandolo, C. M. & Selby, J. H. Ruling out higher-order interference from purity principles. *Entropy* **19**, 253 (2017).

19. Barnum, H., Lee, C. M. & Selby, J. H. Oracles and query lower bounds in generalised probabilistic theories. *Found. Phys.* **48**, 954 (2018).

20. Garner, A. J. P. Interferometric computation beyond quantum theory. *Found. Phys.* **48**, 886 (2018).

21. Lee, C. M. & Selby, J. H. Generalised phase kick-back: the structure of computational algorithms from physical principles. *New J. Phys.* **18**, 033023 (2016).

22. Lee, C. M. & Hoban, M. J. Bounds on the power of proofs and advice in general physical theories. *Proc. R. Soc. A* **472**, 20160076 (2016).

23. Jordan, P., von Neumann, J. & Wigner, E. On an algebraic generalization of the quantum mechanical formalism. *Ann. Math.* **35**, 29–64 (1934).

24. Dakić, B. & Brukner, Č. The classical limit of a physical theory and the dimensionality of space. In *Quantum Theory: Informational Foundations and Foils* (eds. Chiribella, G. & Spekkens, R. W.) (Springer, Dordrecht, 2016).

25. Aaronson, S. Is quantum mechanics an island in theoryspace? Preprint at arXiv:quant-ph/0401062 (2004).

26. De la Torre, G., Masanes, Ll, Short, A. J. & Müller, M. P. Deriving quantum theory from its local structure and reversibility. *Phys. Rev. Lett.* **109**, 090403 (2012).

27. Thompson, G. Normal forms for skew-symmetric matrices and hamiltonian systems with first integrals linear in momenta. *Proc. Am. Math. Soc.* **104**, 910–916 (1988).

28. Graydon, M. A Quaternions and Quantum Theory. Master's thesis, University of Waterloo (2011).

29. Hardy, L. Quantum theory from five reasonable axioms. Preprint at arXiv:quant-ph/0101012 (2001).

30. Dakić, B. & Brukner, Č. Quantum theory and beyond: is entanglement special?. In *Deep Beauty: Understanding the Quantum World through Mathematical Innovation* (ed. Halvorson, H.) (Cambridge University Press, Cambridge, 2011).

31. Masanes, Ll & Müller, M. P. A derivation of quantum theory from physical requirements. *New J. Phys.* **13**, 063001 (2011).

32. Chiribella, G., D'Ariano, G. M. & Perinotti, P. Informational derivation of quantum theory. *Phys. Rev. A* **84**, 012311 (2011).

33. Hardy, L. Reformulating and reconstructing quantum theory. Preprint at arXiv:1104.2066 (2011).

34. Höhn, P. A. Toolbox for reconstructing quantum theory from rules on information acquisition. *Quantum* **1**, 38 (2017).

35. Höhn, P. A. & Wever, A. Quantum theory from questions. *Phys. Rev. A* **95**, 012102 (2017).

36. Höhn, P. A. Quantum theory from rules on information acquisition. *Entropy* **19**, 98 (2017).

37. Goyal, P. Derivation of quantum theory from Feynman's rules. *Phys. Rev. A* **89**, 032120 (2014).

38. Wilce, A. A royal road to quantum theory (or Thereabouts). Preprint at arXiv:1606.09306 (2016).

39. Selby, J. H., Scandolo, C. M. & Coecke, B. Reconstructing quantum theory from diagrammatic postulates. Preprint at arXiv:1802.00367 (2018).

40. Van de Wetering, J. Reconstruction of quantum theory from universal filters. Preprint at arXiv:1801.05798 (2018).

41. Randall, L. *Warped Passages: Unraveling the Mysteries of the Universe's Hidden Dimensions* (Harper Perennial, New York, 2006).

42. Agashe, K. & Pomarol, A. Focus on extra space dimensions. *New J. Phys.* **12**, 075010 (2010).

43. Kaluza, T. Zum Unitätsproblem der Physik. *Akad. Wiss. Berlin* **33**, 966–972 (1921).

44. Klein, O. Quantentheorie und fünfdimensionale Relativitätstheorie. *Z. für Phys. A* **37**, 895–906 (1926).

45. Arkani-Hamed, N., Dimopoulos, S. & Dvali, G. The Hierarchy problem and new dimensions at a millimeter. *Phys. Lett. B* **429**, 263–272 (1998).

46. Antoniadis, I. A possible new dimension at a few TeV. *Phys. Lett. B* **246**, 377–384 (1990).

47. Pawlowski, M. & Winter, A. Hyperbits: the information quasiparticles. *Phys. Rev. A* **85**, 022331 (2012).

48. Araki, H. On a characterization of the state space of quantum mechanics. *Commun. Math. Phys.* **75**, 1–24 (1980).

49. Barnum, H., Graydon, M. A. & Wilce, A. Some nearly quantum theories. *EPTCS* **195**, 59–70 (2015).

50. Barnum, H., Graydon, M. A. & Wilce, A. Composites and categories of Euclidean Jordan algebras. Preprint at arXiv:1606.09331 (2016).

51. Namioka, I. & Phelps, R. R. Tensor products of compact convex sets. *Pac. J. Math.* **31**, 469–480 (1969).

52. Bennett, C. H. Logical reversibility of computation. *IBM J. Res. Dev.* **17**, 525–532 (1973).

53. Coecke, B. & Kissinger, A. *Picturing Quantum Processes – A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge University Press, Cambridge, 2017).

54. Masanes, Ll, Müller, M. P., Augusiak, R. & Pérez-García, D. Existence of an information unit as a postulate of quantum theory. *Proc. Natl Acad. Sci. USA* **110**, 16373 (2013).

55. Masanes, Ll, Müller, M. P., Augusiak, R. & Pérez-García, D. Entanglement and the three-dimensionality of the Bloch ball. *J. Math. Phys.* **55**, 122203 (2014).

56. Müller, M. P. & Masanes, Ll Three-dimensionality of space and the quantum bit: an information-theoretic approach. *New J. Phys.* **15**, 053040 (2013).

57. Weinberg, S. Testing quantum mechanics. *Ann. Phys.* **194**, 336–386 (1989).

58. Gisin, N. Weinberg's nonlinear quantum mechanics and supraluminal communications. *Phys. Lett. A* **143**, 1–2 (1990).

59. Galley, T. D. & Masanes, Ll Classification of all alternatives to the Born rule in terms of informational properties. *Quantum* **1**, 15 (2017).

60. Galley, T. D. & Masanes, Ll. Impossibility of mixed-state purification in any alternative to the Born rule. Preprint at arXiv:1801.06414 (2018).

61. Jacobson, R. Thermodynamics of spacetime: the einstein equation of state. *Phys. Rev. Lett.* **75**, 1260 (1995).

62. Ryu, S. & Takayanagi, T. Holographic derivation of entanglement entropy from the anti-de Sitter space/conformal field theory correspondence. *Phys. Rev. Lett.* **96**, 181602 (2006).

63. Höhn, P. A. Reflections on the information paradigm in quantum and gravitational physics. *J. Phys.* **880**, 012044 (2017).

64. Höhn, P. A. & Müller, M. P. An operational approach to spacetime symmetries: Lorentz transformations from quantum communication. *New J. Phys.* **18**, 063026 (2016).

65. Müller, M. P., Carrozza, S. & Höhn, P. Is the local linearity of space-time inherited from the linearity of probabilities? *J. Phys. A* **50**, 054003 (2017).

66. Garner, A. J. P., Müller, M. P. & Dahlsten, O. C. O. The complex and quaternionic quantum bit from relativity of simultaneity on an interferometer. *Proc. R. Soc. A* **473**, 20170596 (2017).

67. Aaronson, S. NP-complete problems and physical reality. *ACM SIGACT News* **36**, 30–52 (2005).

68. Simon, B. *Representations of finite and compact groups* (American Mathematical Society, Providence, RI, 1996).