Computation in a general physical setting

Ciarán M Gilligan-Lee

Department of Physics and Astronomy, University College London, United Kingdom
E-mail: ciaran.lee@ucl.ac.uk

Received 23 March 2021, revised 28 June 2021
Accepted for publication 23 August 2021
Published 8 September 2021

Abstract

The computational abilities of theories within the generalised probabilistic theory framework has been the subject of much recent study. Such investigations aim to gain an understanding of the possible connections between physical principles and computation. Moreover, comparing and contrasting the computational properties of quantum theory with other operationally-sensible theories could shed light on the strengths and limitations of quantum computation. This paper reviews and extends some of these results, deriving new bounds on the computational ability of theories satisfying \( n \)-local tomography, and theories in which states are represented as generalised superpositions. It moreover provides a refined version of the conjecture that a quantum computer can simulate the computation in any theory within a certain sub-class of generalised probabilistic theories with at most polynomial overhead. The paper ends by describing an important relation between this conjecture and delegated computation, similar to the relation between quantum non-locality and device-independent cryptography.

Keywords: quantum computation, generalised probabilistic theories, quantum foundations

1. Introduction

Quantum theory is a strange beast; its predictions have been verified to unprecedented accuracy, yet the standard quantum formalism—in which quantum states are represented by positive semi-definite operators acting on an underlying complex Hilbert space—is as abstract as its predictions are accurate. Despite being universally accepted among physicists as a tool for calculating the probabilities of possible experimental outcomes, the standard language of complex Hilbert spaces and positive operators lacks direct physical or operational significance. As Asher
Peres [1] famously put it: ‘quantum phenomena do not occur in a Hilbert space. They occur in a laboratory’.

Taking inspiration from Einstein’s operational formulation of special relativity, in which reference frames are operationally defined using clocks and rods, researchers have begun to study quantum theory from an operational perspective. Researchers have formulated quantum theory within an operational framework of generalised probabilistic theories—which generalise the probabilistic formalism of quantum theory. Remarkably, researchers have derived the structure of finite dimensional quantum theory [2–7] within this framework from operationally-framed physical principles—similar to Einstein’s derivation of the Lorentz transformations from two physical principles: the constancy of the speed of light and the principle of relativity. Researchers have even used this framework to rule out certain a priori reasonable types of post-quantum physics [8, 9].

A remarkable feature of this operational framework is that it provides examples of theories which differ from quantum theory, yet still make good operational sense. An obvious example is classical probability theory, which can be used to calculate probabilities for classical situations such as tossing a coin, or conducting an experiment in the regime of Newtonian physics. More exotic examples include Spekkens’ toy theory [17, 18], a construction colloquially known as ‘Boxworld’ [6, 19], which achieves the largest possible violation of the CHSH inequality [20] consistent with the no-signalling principle, and the theory referred to by the authors as ‘Witworld’ [21], which exhibits post-quantum steering.

The existence of such alternate theories allows for an investigation of the structural or information-theoretic properties of theories where different physical principles may hold. Such an investigation could provide a deep understanding of the possible connections between physical principles and information-theoretic advantages in a manner not wedded to the mathematical formulation of a specific theory [4, 6, 22, 23]. This forms part of the broader research program of generalised probabilistic theories, which, in the words of Barnum, Müller, and Ududec [24], aims to ‘analyse the structure of physics—that is, the way that the different parts of physics fit together—by rigorously assessing the consequences of changing some of its parts’. Moreover, by comparing and contrasting the information-theoretic properties of quantum theory with other theories in this framework, one may shed light on the strengths and limitations of quantum information processing and quantum computing in particular.

The study of computation within the generalised probabilistic theory framework was initiated in the work of Barrett [6]. An intriguing aspect of that work resulted in a conjecture concerning the power of quantum computing relative to other theories within this operational framework. Specifically, it was conjectured that quantum theory may be able to simulate the computation in any generalised probabilistic theory with at most polynomial overhead. This conjecture has spurred much recent work [23, 25–30]. Reference [25] showed that in any theory satisfying the principle of tomographic locality, which informally states that local measurements suffice for tomography, the upper bound on efficient computation is the same as the best known upper bound on efficient quantum computation. Additionally, this bound holds regardless of whether the principle of causality, which roughly states that there is no signalling from the future, is satisfied. Moreover, reference [26] showed that by slightly altering the definition of a generalised theory one can construct a theory in this modified framework that satisfies both tomographic locality and causality which achieves this upper bound, and hence can simulate any quantum computation efficiently. The current paper reviews and extends some of these results and concludes with a more nuanced version of the conjecture originally made in

1 For further examples, see references [10–16].
2. Generalised theories

A fundamental goal of any physical theory is to provide a consistent explanation of experimental data. This constitutes the core idea underlying the framework of generalised probabilistic theories, where the primitive notions are operational\(^2\) in nature. Indeed, as any candidate physical theory will eventually be experimentally investigated, it should have an operational description in terms of these experiments. We will work in the circuit framework for generalised probabilistic theories developed by Hardy in [3] and Chiribella, D’Ariano, and Perinotti in [4]. The circuit framework takes inspiration from the categorical approach to quantum theory, introduced by Abramsky and Coecke [31, 32], in that it heavily emphasises compositionality. This compositional viewpoint also has advantages in standard quantum information, see for example reference [33]. In this manner the circuit framework is different to the convex sets framework, an alternatively approach to generalised probabilistic theories, which was presented in [2, 6]—although both are similar in spirit.

Informally, a theory in this framework specifies a set of laboratory devices that can be connected together in different ways and assigns probabilities to different experimental outcomes [3, 4, 6, 8]. Devices are equipped with a number of input and output ports, together with a classical pointer that can take various distinct positions. One constructs experiments by connecting the output ports of some number of devices to the input ports of others until there are no unconnected ports remaining, ensuring along the way that no cycles have been formed. When these devices are used in an experiment, the classical pointer comes to rest in a specific position, denoting the experimental outcome. Intuitively, one can think of a physical system as passing between a devices ports. These systems come in different types, denoted \(A, B, C, \ldots\). When composing devices in sequence and parallel to form experiments, the types of connected ports must match. Moreover, the set of devices is closed under sequential and parallel composition.

In the current framework, experiments—closed circuits of devices with matching types and no cycles—correspond to probabilities over the possible classical pointer positions of each constituent device in the experiment. Moreover, if a closed circuit consists of multiple disconnected circuits, then the probability assignment factorises over this decomposition. Devices with the same input and output types yielding the same probabilities in all possible experiments are identified. The equivalence class of devices with no input ports are called states, no output ports effects, and both input and output ports transformations. For input type \(A\) and output type \(B\), these sets will be denoted \(\text{St}(B)\), \(\text{Eff}(A)\), and \(\text{Transf}(A, B)\).

The correspondence between closed circuits and probabilities induces a linear structure that will be essential in the remainder of the paper. Note that states, effects, and transformations can be thought of as functions from the set of effects, states, and circuit fragments of matching type\(^3\) to the interval \([0, 1]\) [4, 25]. As one can take linear combinations of functions, the set of states, effects, and transformations of a given type each respectively can be seen to generate a real vector space, in which the set of states, effects, and transformations are embedded. Note that this embedding in general results in a subset rather than a subspace. Moreover, the sets

\(^2\)Note that operationalism as a philosophical viewpoint, in which one asserts that there is no reality beyond laboratory device settings and outcomes, is not being espoused here. One should merely view the approach taken here as an operational methodology aimed at gaining insight into certain structural properties of physical theories.

\(^3\)Circuit fragments of matching type correspond to those whose unconnected ports are of the same type as the ports of the corresponding transformation.
of effects and transformations act linearly on the vector space generated by the set of states of the appropriate type. The vector space generated by $\text{St}(B)$, $\text{Eff}(A)$, and $\text{Transf}(A, B)$ will be denoted $V_B$, $V_A$, and $V_{AB}$ respectively.

In this work the standard assumption that the above vector spaces are all finite dimensional will be made. This corresponds to the operational statement that there exists a finite set of fiducial measurements whose statistics are sufficient for the tomography of arbitrary states. This implies that the vector space of state and effects are dual to one another [4, 25]. In quantum theory, such a collection of measurements is called informationally complete [35, 36]. The correspondence between states and real vectors in the space they generate can be chosen in manifold different—yet statistically equivalent—ways [37]. Indeed, one can always choose to represent states as real vectors whose entries correspond to the outcomes of fiducial measurements. Given a finite set of fiducial measurements, they can be combined into a single measurement consisting of a coin flip to select a measurement from the set, followed by an application of said measurement [38]. Hence one can always choose to represent a given state in such a theory as a real vector whose entries correspond to the probabilities of specific outcomes of such a single fiducial measurement.

A given measurement corresponds to a set of effects $\{e^r\}$ labelled by the position of the classical pointer $r$ attached to the measurement device. The probability of preparing some state $s$, observing that transformation $T_k$ has been applied, and observing outcome $r$ is (suppressing system types for readability) given by:

$$e^r(T_k s) = P(r, k, s).$$

Hence the probabilities for a given circuit are calculated by performing matrix multiplication of the vectors, matrices, and dual vectors associated with the states, transformations and measurement outcomes in the given closed circuit/experiment [4, 25].

As mentioned previously, finite-dimensional quantum theory is an example of a generalised probabilistic theory. A quantum system is associated with a complex Hilbert space, with the system type given by the dimension of the Hilbert space. States and effects are associated with positive operators, transformations with trace non-increasing completely positive maps. A device with no input ports corresponds to what is sometimes called a ‘random source of quantum states’, and is associated with positive operators $\{\rho_r\}$, with $r$ denoting the possible positions of the classical pointer, such that $\sum_r \text{Tr}(\rho_r) = 1$. When device is used in an experiment, the probability that the classical pointer takes position $r$ is given by $\text{Tr}(\rho_r)$, and the quantum state prepared, conditioned on the pointer reading being $r$, is the normalised operator $\rho_r/\text{Tr}(\rho_r)$. A device with no output ports is associated with a positive operator-valued measurement, that is a set of positive operators $\{E_i\}$ satisfying $\sum_i E_i = \mathbb{1}$. A device with both input and output ports is associated with a quantum instrument, that is a set of trace non-increasing completely positive maps, one for each value of the pointer reading $r$, that sum to a trace-preserving map. Given such associations, the usual rules of quantum theory allow for calculation of probabilities for any circuit outcome.

The central physical principles needed in formulation of the main results of this paper will now be formally stated.

**Definition 1 (Causality [4]).** A theory is causal if there exists a unique deterministic effect (the device corresponding to such an effect only has a single pointer position) for all system types.

---

4 See reference [34] for a nuanced discussion of this point.
Mathematically, the principle of causality is equivalent to the statement: ‘probabilities of present experiments are independent of future measurement choices’. Hence it captures the intuitive notion that there should be ‘no-signalling from the future’ [4]. Additionally, the causality principle implies the standard ‘no-signalling in space’ principle [4, 39]. Moreover, the unique deterministic effect allows one to define a notion of marginalisation for multi-partite states. Causality need not be satisfied by all generalised probabilistic theories, indeed reference [14] has explicitly constructed a theory which does not satisfy the principle of causality. Quantum theory is causal, the unique deterministic effect being given by the identity matrix $I_d$ for a $d$-level system.

**Definition 2 (Tomographic locality for states [2, 4, 6]).** A theory satisfies tomographic locality for states if the state of every composite system is uniquely specified by the statistics of measurements performed locally on each component system.

Given two distinct composite states of the same type, tomographic locality implies the existence of local effects which, when applied to each individual system, give different probabilities on each state. One can generalise this principle to involve transformations as follows.

**Definition 3 (Tomographic locality for transformations [25]).** A theory satisfies tomographic locality for transformations if every transformation is uniquely characterised by local process tomography, that is, by inputting local states and performing local measurements on each component system.

Violation of tomographic locality (either for states or transformations) corresponds to the existence of ‘global degrees of freedom’ that are not accessible at the level of local measurements. A consequence of tomographic locality is that for a transformation with input type $AB$ and output type $CD$, the corresponding real vector space has the following decomposition [4, 6, 25]:

$$V_{AB}^{CD} \cong V_A \otimes V_B \otimes V_C \otimes V_D,$$

where $\otimes$ here denotes the ordinary vector space tensor product. In particular, tomographic locality for states implies that for a bipartite state of type $AC$, the corresponding vector space decomposes as $V_{AC} \cong V_A \otimes V_C$. Furthermore, a transformation $T_x \in \text{Transf}(A, B)$ is completely specified by its action on $\text{St}(A)$ [4]. Quantum theory satisfies both tomographic locality for states and transformations [4]. For the case of states, one can indeed check that the real vector spaces of Hermitian operators satisfies $V_{AB} \cong V_A \otimes V_B$.

**Definition 4 (n-local tomography for states [40]).** A theory is $n$-locally tomographic if the state of every composite system is uniquely specified by the statistics of measurements performed on each single system, each pair of systems, each triple of systems, . . . , each $n$-tuple of systems belonging to the composite.

As for tomographic locality, one can define an analogous principle of $n$-local tomography for transformations. Given two distinct composite states of the same type, $n$ local tomography implies the existence of $n$-local effects, consisting of conical combinations of effects that act on at most $n$-systems, which give different probabilities on each state. As $n$-local theories violate tomographic locality for $n > 1$, there is a global degree of freedom which cannot be accessed by local measurements. However, as we’ll see in section 3, such global parameters do not grow ‘too fast’ with increasing system size.

---

5 See reference [34] for a slight subtlety regarding the definition of tomographic locality in the circuit approach to generalised probabilistic theories as opposed to the convex sets approach.
Quantum theory defined over real—rather than complex—Hilbert spaces supplies an example of a theory that is \( n = 2 \)-local, or bi-local, and hence does not satisfy tomographic locality. In real quantum theory, states and effects correspond to real symmetric matrices (which satisfy the same set of constraints as the standard complex quantum case discussed above), and transformations to completely positive maps that preserve the set of real symmetric matrices. We will now provide an example of two distinct transformations between two two-dimensional systems (sometimes referred to as rebit’s [13]), which cannot be locally distinguished in the theory and so provides a violation of tomographic locality. Consider

\[
T_1(\rho) = \frac{1}{2}\rho + \frac{1}{2}Y\rho Y, \quad \text{and} \quad T_2(\rho) = \frac{1}{2}I \cdot \text{Tr}(\rho),
\]

where \( Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \) is the Pauli \( Y \) matrix and \( T_2 \) is a measure and prepare transformation that traces out the input state and prepares the maximally mixed state, \( \frac{1}{2}I \). Clearly the transformation \( Y \cdot Y \) is allowed in the theory as:

\[
Y\rho Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} a & b \\ b & 1-a \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} = \begin{pmatrix} 1-a & -b \\ -b & a \end{pmatrix}, \quad \forall a, b \in \mathbb{R}.
\]

It is easy to see that \( T_1(\rho) = T_2(\rho) \) for all real symmetric matrices \( \rho \) with trace one. Hence, one cannot distinguish these two transformations by inputting a single rebit and performing a single measurement on the output state. To distinguish \( T_1 \) from \( T_2 \), one has to evaluate them on one half of the Bell state \( |\phi^+\rangle \langle \phi^+| \) and perform a joint measurement of the two output systems. One has

\[
T_1 \otimes I \left( |\phi^+\rangle \langle \phi^+| \right) = \frac{1}{2} |\phi^+\rangle \langle \phi^+| + \frac{1}{2} |\psi^-\rangle \langle \psi^-|, \quad \& \quad T_2 \otimes I \left( |\phi^+\rangle \langle \phi^+| \right) = \frac{1}{2}I \otimes I.
\]

Performing the two-outcome measurement \( \{ |\phi^+\rangle \langle \phi^+| + |\psi^-\rangle \langle \psi^-|, |\phi^-\rangle \langle \phi^-| + |\psi^+\rangle \langle \psi^+| \} \), where \( |\phi^\pm\rangle = \{ |00\rangle \pm |11\rangle \} / \sqrt{2} \) and \( |\psi^\pm\rangle = \{ |01\rangle \pm |10\rangle \} / \sqrt{2} \), distinguishes these two states. Note that as

\[
(T_1 \otimes I - T_2 \otimes I)(|\phi^+\rangle \langle \phi^+|) = \frac{1}{4} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \otimes \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (2.1)
\]

we have \( \text{Tr}((T_1 \otimes I - T_2 \otimes I)(E \otimes F)) = 0 \) for all real symmetric matrices \( E, F \). Hence one cannot distinguish these two states with local measurements. Moreover, one can think of equation (2.1) as a global degree of freedom not accessible to local observers. It was shown by Hardy and Wooters in [40] that one only ever needs to perform joint measurements between at most two subsystems to distinguish any two states in real quantum theory.

3. Computation

The class of problems a quantum computer can efficiently solve, and whose answer can be stated as either ‘accept’ or ‘reject’, is denoted \( \text{BQP} \). Since the very beginning of Quantum Computation, much research has involved placing upper bounds on this class. Put another way, much research in quantum computing have been concerned with how large this class is. At present, the best known upper bound is that \( \text{BQP} \subseteq \text{AWPP} \), where \( \text{AWPP} \) is a classical complexity class—known to be contained in \( \text{PP} \), hence \( \text{PSPACE} \)—with a slightly obscure definition involving ‘gap’ functions for non-deterministic Turing machines. We will provide
a much more intuitive definition\(^6\) in section 3.3 which first originated in [26]. Phrased alternatively: a quantum computer cannot efficiently solve any problem outside the class \textit{AWPP}, but it is not known whether it can solve every problem contained within \textit{AWPP}. Indeed, it is believed quite likely that \textit{BQP} is a strict subset of \textit{AWPP} [26].

In the following subsections, we present two different—yet equivalent—models of computation in the generalised probabilistic theory framework. The first one generalises the standard classical and quantum circuit model, and the second involves a modification of the probabilistic Turing machine model. Additionally, we review and expand previous results involving an upper bound on the power of efficient computation in a certain class of generalised probabilistic theories which first appeared in [25]. Finally, we discuss how oracles are defined in generalised theories and concludes by reviewing and extending certain oracle separation results due to [28, 30].

### 3.1. Circuit model

The framework introduced in section 2, in which devices can be connected together in sequence and parallel to form circuits, suggests a natural model of computation which generalises the classical and quantum circuit model of computation. In order to formally define an efficient circuit model of computation in theories belonging to the framework introduced in section 2, the notion of a (polynomially sized) uniform circuit family is needed, together with a condition for a circuit to accept or reject an input. We now present such a definition, providing an intuitive explanation of each point after the formal definition. A polynomially sized uniform circuit family is a set of closed circuits \(\{C_x\}\) such that:

(a) There is a finite gate\(^7\) set \(G\), consisting of devices, such that each circuit in the family is built from elements of \(G\).

(b) The number of gates in the circuit \(C_x\) is bounded by a polynomial in \(|x|\).

(c) There is a Turing machine that, acting on input \(x = x_1x_2\ldots x_n\), outputs a classical description of \(C_x\) in time bounded by a polynomial in \(|x|\).

(d) For each type of system, there is a fixed choice of basis, relative to which transformations are associated with matrices. Given the matrix \(M\) representing (a particular outcome of) a gate in \(G\), a Turing machine can output a matrix \(\tilde{M}\) with rational entries, such that \(|(M - \tilde{M})_{ij}| \leq \epsilon\), in time polynomial in \(\log(1/\epsilon)\).

Regarding item 1, for a specific theory in our framework it may not be the case that gates acting only on bipartite and single systems are universal for computation, as is the case for quantum computation. Thus for any \(m, k\), a circuit might involve gates with \(m\) input systems and \(k\) output systems. In general, it may be that no finite gate set is universal for computation. Nonetheless, we demand as a requirement of uniformity that any uniform circuit family is built from elements of a finite gate set\(^8\). A consequence is that the number of distinct system types appearing in a given uniform circuit family is also finite.

Regarding item 2 and 3, as we saw in section 2, not every generalised probabilistic theory satisfies the principle of causality, in which case a circuit does not have a ‘preferred’ direction. Hence there is no reason to assume—as is standardly done for classical and quantum circuits—that a circuit must have the form of a number of gates acting on some input, where

---

\(^6\) See references [41, 42] for another definition of \textit{AWPP} which physicists may also find intuitive.

\(^7\) When discussing computation, the terms ‘device’ and ‘gate’ will be used interchangeably.

\(^8\) One may alternatively consider a uniformity condition in which the number of allowed gates increases with growing circuit size, as is the case in [43, section 3.3 A].
the input preparation encodes the problem instance. Instead, we permit the entire circuit to encode the problem instance, defining a circuit family as a set \( \{ C_x \} \) with the stipulation that each circuit is indexed by a classical string \( x = x_1 x_2 \ldots x_n \). A circuit family is polynomial-size if the number of gates is bounded from above\(^9\) by a polynomial in \(|x|\). Note that there are formulations of quantum circuits in which the entire circuit encodes the problem instance rather than just the input states, see section 2.2 of reference [44] for a specific example.

Regarding item 4, the final requirement for a circuit family to be uniform manifests as a constraint on the entries of the matrices representing the transformations that appear in the finite gate set. If such a constraint is not made, it may be possible to smuggle hard to compute quantities into the computation through the matrix entries. Note that such a constraint is required even in the classical case. Indeed, consider a coin whose probability of landing heads up is equal to Chaitin’s constant [45], which, informally, represents the probability that a given Turing machine will halt on a randomly chosen input. By flipping such a coin multiple times, one could extract the first \( N \) bits of this probability. Using such an approximation one could solve the halting problem for all inputs of size up to \( N \) in some fixed amount of time. This is believed not to be possible, hence there should exist some constraint on which probabilities are accessible computationally. We can motivate such a constraint by recalling that gates correspond to operational devices; an experimenter with access to devices governed by some generalised probabilistic theory may only be able to characterise them tomographically to finite precision—and the features of a probabilistic theory should not be sensitive to precision issues which are inaccessible to experiment. We formalise this constraint as follows: there must exist some fixed choice of basis for \( V_A \) for each system \( A \), such that a Turing machine can efficiently compute approximations to the entries of the matrices relative to these bases. We require that for any matrix entry \((M)_{ij}\), and any \( \epsilon \), a Turing machine can output a rational number, within \( \epsilon \) of \((M)_{ij}\), in time bounded by a polynomial in \( \log(\frac{1}{\epsilon}) \).

Denoting the string of observed outcomes by \( z \), the final output of the computation will be given by a function \( a(z) \in \{0, 1\} \). We say that a run of the experiment accepts an input string \( x \) if the outcome string \( z \) of the circuit \( C_x \) satisfies \( a(z) = 0 \). We require the existence of a Turing machine that computes \( a \) in time polynomial in the length of the input \(|x|\). This final requirement is needed to ensure that the calculation of whether a given circuit outcome accepts or rejects should not itself be able to solve a computationally hard problem. The probability that a computation accepts the input string \( x \) is thus

\[
P_x(\text{accept}) = \sum_{z | a(z) = 0} P(z),
\]

where the sum ranges over all possible outcome strings \( z \) of the circuit \( C_x \) for which \( a(z) = 0 \).

In the quantum case, one usually says a given computation accepts its input if a measurement of the first outcome qubit in the computational basis yields \( |0\rangle \). This is due to the fact that the application of any non-deterministic quantum operation at an intermediate stage in the computation can always without loss of generality be replaced by a unitary operation (potentially acting on some extra auxiliary qubits) followed by a projective measurement at the end of the computation. Hence, the only outcomes, i.e. non-trivial classical pointer positions, generated by the computation appear on measurements performed at the end of the circuit. In general it is not the case that one can postpone non-deterministic operations (that is, operations represented by a device with more than a single classical pointer position) to the end of a circuit.

---

\(^9\) By padding with identity transformations, one can always ensure that the number of gates is exactly specified by a polynomial in \(|x|\).
in an arbitrary generalised probabilistic theory.\footnote{If the theory under consideration satisfies the purification principle of [4], then non-deterministic measurements can indeed be postponed to the end of a computation as in such theories any non-deterministic operation can be ‘dilated’ to a reversible operation followed by a measurement. In fact, the purification principle is equivalent to the existence of such a reversible dilation for every operation in the theory [4].} Hence the need for the generalised acceptance condition outlined above.

Given the notion of a uniform circuit family and our generalised acceptance condition, we can now formally define the class of problems that can be efficiently solved in an arbitrary generalised probabilistic theory.

**Definition 5 (Class of efficient computation in a generalised probabilistic theory [25]).** For a theory $G$, a language $L$ is in the class $BGP$ if there exists a poly-sized uniform family of circuits in $G$, and an efficient acceptor, such that

(a) $x \in L$ is accepted with probability at least $\frac{2}{3}$.

(b) $x \notin L$ is accepted with probability at most $\frac{1}{3}$.

The constants in the above definition can be chosen arbitrarily as long as they are bounded away from a half by an inverse polynomial in $x$ [25].

The following theorem was proved in [25]:

**Theorem 6 (Theorem 3.4.1 in [25]).** For any generalised probabilistic theory $G$ satisfying tomographic locality, the following holds:

$$BGP \subseteq AWPP \subseteq P\text{P} \subseteq \text{PSPACE}.$$

Note that the principle of causality was not required to derive the above bound. Once the appropriate definitions are in place, the proof of theorem 6 is a fairly straightforward extension of similar proofs in the quantum case [46]. We will now give an intuitive overview of how the operational assumptions underlying the generalised probabilistic theory framework together with tomographic locality allow for the derivation of the bound from theorem 6.

First, the proof relies on the fact that transformations in a generalised probabilistic theory are linear, and hence have a matrix representation. As we saw in section 2 this linear structure arose from the intuitive requirement that a physical theory should be able to give probabilistic predictions about the occurrence of possible outcomes. Second, a further requirement of the proof is the ability to compute efficiently the entries in the matrices representing the transformations applied in parallel in a specific circuit. Recall from section 2, that in a tomographically local theory the matrix corresponding to transformations applied in parallel can be easily calculated by taking the tensor product of the matrices representing each individual transformation. As the matrices for individual gates satisfy the uniformity condition, the elements of tensors products of these matrices do as well. By foliating the circuit so that only a single non-trivial gate acts in any given foliation, one sees that the outcome probability of the circuit corresponds to the multiplication of matrices consisting of tensor products of these gates with identity transformations on systems on which they do not act [25]. Generically, this is not the case without tomographic locality—as the tensor product structure may not hold.

If a transformation from $A$ to $B$ acts only on a subsystem of $AC$, there might be no simple relation between the linear map $\text{St}(AC) \rightarrow \text{St}(BC)$ and the arising from the action of the transformation when it is applied to system $A$ on its own, or indeed to a joint system $AD$, for some other system $D$. There may thus be no efficient way of computing matrix elements corresponding to a transformation considered as part of a circuit of arbitrary size. Recall that one can think of a violation of tomographic locality as akin to the existence of ‘global degrees
of freedom’ not accessible via local measurements. As the uniformity condition only imposes constraints on the matrices associated to gates from the finite gate set $\mathcal{G}$, if there doesn’t exist an efficient procedure for computing the matrices associated to the parallel compositions of gates from $\mathcal{G}$ one could in principle encode answers to computationally hard problems in these global degrees of freedom. While tomographic locality implies such an efficient procedure, it may not be the simplest principle to do so.

Real Hilbert space quantum theory, discussed at the end of section 2, provides an example of a theory without tomographic locality for which the bound of theorem 6 still holds, since there is an efficient way of calculating the relevant matrix entries. One can in fact show that any theory which satisfies $n$-local tomography, for $n$ a fixed constant, also satisfies the bound of theorem 6.

In our original set up, our uniformity assumption ensured that the entries of each matrix from our finite gate set could be efficiently approximated by a Turing machine, i.e. a ‘classical’ computer. One could imagine the classical computer being fed the probabilities of fiducial measurement outcomes, from which it efficiently calculates the matrix corresponding to this transformation in some fixed basis. This statement just sums up the fact that once one has collected all the necessary statistics, one can ‘easily’ output the state (or transformation) they describe. Without such a commitment, one could not do tomography. Hence to output an approximation of the matrix entry of some transformation, a classical computer has to hold in its memory the probabilities of all fiducial measurement outcomes. From this it can easily approximate (to some degree of error) the required matrix entry.

Now, if we assume tomographic locality then the classical computer only needs to store a finite number of these fiducial probabilities. These are just the fiducial outcomes for the system on which the non-identity transformation acts, as the matrix entries for the tensor product of this transformation with identities are either zero, or just the entries of the non-identity transformation itself (from uniformity of the circuit, the computer can efficiently compute on which system the non-identity transformation acts). In the case of $n$-local tomography, for $n$ a fixed constant, the number of measurements needed to specify a given transformation grows with system size, but it doesn’t grow ‘too fast’. For a transformation acting on $N$ systems, the number of measurements required is

$$k \left( \frac{N!}{n!(N-n)!} \right),$$

where $k$ is the largest integer such that the outcome of $k$ fiducial measurements is sufficient to determine any state of a given system. Note that as only finitely many systems appear in a given circuit family (item 1 from the statement of uniformity), $k$ is a constant. To leading order, equation (3.1) goes as $cN^n$, with $c$ and $n$ constant. Hence the classical computer needs to store a polynomially growing set of fiducial measurement outcomes; it needs a poly-size memory. However, this machine only requires at most poly-time to run, acting in some sense as a transducer which writes the description of the matrix on an output tape. Hence a standard poly-size, poly-time Turing machine suffices.

Given the above argument, the remainder of the proof of theorem 6 in [25] goes through verbatim. We have thus provided a sketch proof of the following theorem.

**New theorem (Extends theorem 3.4.1 in [25]).** For any generalised probabilistic theory $\mathcal{G}$ satisfying $n$-local tomography, for $n$ a fixed constant, the following bound holds:

$$\text{BGP} \subseteq \text{AWPP} \subseteq \text{PP} \subseteq \text{PSPACE}.$$
If a theory does not even satisfy \( n \)-local tomography for any fixed \( n \), then the number of fiducial outcomes needed to specify a transformation could grow exponentially in system size. A poly-size Turing machine would not be able to store all of these in its memory, hence it appears the bound will not hold in these situations.

### 3.2. Oracles

In quantum computing, oracles are a vital component of many of the known computational speed-ups over classical computing [47–49]. In general, an oracle provides access to some function \( f : \{1, \ldots, N\} \to \{0, 1\} \) and one is usually concerned with the number of queries to this oracle needed to solve a certain problem. In quantum computing an oracle corresponds to a controlled unitary transformation \( U_f \) (in fact, to a family of controlled unitaries—one for each problem size) which acts on the computational basis \( \{|x\}\} \) as

\[
U_f \left| x \right\rangle \left| y \right\rangle = \left| x \right\rangle \left| y \oplus f(x) \right\rangle.
\]

By performing measurements on the target system, represented above by the \( |y\rangle \) state, one can learn about the value of \( f \) on specific inputs.

However, a generic generalised probabilistic theory does not have sufficient structure to define such an oracle [25]. Without one, it is difficult to compare the computational abilities of different theories and hence to assess how computational power depends on different physical principles. In reference [25] a rudimentary oracle model was defined for all theories satisfying the principle of causality. It was shown that, relative to this oracle model, the class \( \text{NP} \) was not contained in any theory satisfying tomographic locality and causality. A major drawback of this model however, was the fact that it cannot be queried in superposition—that is, in a non-classical manner. However, it was shown in reference [23] that in theories satisfying certain natural physical principles reverser controlled transformations exist and provide a natural generalisation of quantum oracles.

Using this oracle model, reference [30] proved that the best achievable lower bound on the number of quantum queries needed to solve certain query problems is not optimal in the space of all generalised probabilistic theories. An example of such a query problem is \textsc{parity}, a generalisation of Deutsch’s problem [49] which asks for the parity of a function \( f : \{1, \ldots, N\} \to \{0, 1\} \)—that is, the value \( f(1) \oplus \cdots \oplus f(N) \mod 2 \)—where \( N \) is taken to be a constant. On a classical computer, \( N \) queries are needed to solve this problem, but \( \lceil N/2 \rceil \) queries to a quantum oracle suffice to determine the parity. In fact, \( \lceil N/2 \rceil \) queries to a quantum oracle is optimal; a quantum computer cannot determine the parity using fewer queries [50]. However, if a theory exhibits certain type of post-quantum, or higher-order [7, 29, 51–53], interference—which we discuss in more detail below—then a mathematically achievable lower bound to the number of queries needed to determine the parity is 1 [30]—although the physical principles required to reach this bound is unknown in general.

Related work in reference [28] derived Grover’s lower bound to the search problem from simple physical principles. The search problem asks one to find a certain ‘marked item’ from among a collection of items in an unordered database. The only access to the database is through an oracle; when asked if item \( i \) is the marked one, the oracle outputs ‘yes’ or ‘no’. The figure of merit in this problem is how the minimum number of queries required to find the marked item scales with the size of the database. It was shown that, asymptotically,

\[11\] Namely: causality, purification, strong symmetry, and the requirement that the parallel composition of pure states is also a pure state. See [23] for more information.

\[12\] The same principles required to define oracles in generalised theories mentioned in footnote 11.
post-quantum, or higher-order, interference does not provide an advantage over quantum theory in this case, as the number of queries needed to find the marked item in both cases scales as the square root of the database size. For increasing database size $N$, the lower bound to the number of queries scales as $\sqrt{N}$ in both cases. Hence, post-quantum interference is not, in general, a resource for post-quantum computer. Although post-quantum interference may provide a constant speed-up for certain problems, this advantage does not lead to improved scaling with problem size.

The above results were originally derived from the same collection of physical principles (which are outlined in footnote 11). However, those principles were only used to derive a specific, and physically intuitive, representation for states in those theories from first principles. Once the representation is specified, the above results can be proven directly, without reference to the physical principles. Moreover, while those principles are sufficient to derive such a representation, they do not appear to be necessary [54]. Moreover, there is general interest in theories with such representations [54–56]. Hence it is interesting to understand computational consequences of the representation alone, rather than computational consequences of sufficient set of physical principles that imply it.

To describe the abovementioned representation, we first need to define higher-order interference. The definition of higher-order interference that we present here takes its motivation from the set-up of multi-slit interference experiments. In such experiments a particle (a photon or electron, say) passes through slits in a physical barrier. By blocking some of the slits and repeating the experiment many times, one can build up an interference pattern on a screen placed behind the physical barrier. Informally, a theory has ‘$n$th order interference’ if one can generate interference patterns in an $n$-slit experiment which cannot be created in any experiment with only $m$-slits, for all $m < n$.

It was first shown by Sorkin [51, 57] that—at least for ideal experiments [58]—quantum theory is limited to the $n = 2$ case. That is, the interference pattern created in a three—or more—slit experiment can be written in terms of the two and one slit interference patterns obtained by blocking some of the slits.

Given $N$ slits, labelled $1, \ldots, N$, these transformations will be denoted $P_I$, where $I \subseteq \{1, \ldots, N\}$ corresponds to the subset of slits which are not closed. In general one expects that $P_IP_J = P_{I \cup J}$, as only those slits belonging to both $I$ and $J$ will not be closed by either $P_I$ or $P_J$. Thus we think of these transformations as corresponding to projectors (i.e. idempotent transformations $P_IP_I = P_I$). Instead of working directly with these physical projectors, it is mathematically convenient to work with the (generally) unphysical transformations corresponding to projecting onto the ‘coherences’ of a state. Consider the example of a qutrit in quantum theory, the projector $P_{\{0,1\}}$ projects onto a two dimensional subspace:

$$P_{\{0,1\}} :: \begin{pmatrix} \rho_{00} & \rho_{01} & \rho_{02} \\ \rho_{10} & \rho_{11} & \rho_{12} \\ \rho_{20} & \rho_{21} & \rho_{22} \end{pmatrix} \mapsto \begin{pmatrix} \rho_{00} & \rho_{01} & 0 \\ \rho_{10} & \rho_{11} & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

while the coherence-projector $\omega_{\{0,1\}}$ projects only onto the coherences in that two dimensional subspace:

$$\omega_{\{0,1\}} :: \begin{pmatrix} \rho_{00} & \rho_{01} & \rho_{02} \\ \rho_{10} & \rho_{11} & \rho_{12} \\ \rho_{20} & \rho_{21} & \rho_{22} \end{pmatrix} \mapsto \begin{pmatrix} 0 & \rho_{01} & 0 \\ \rho_{10} & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

That is, $\omega_{\{0,1\}}$ corresponds to the linear combination of projectors: $P_{\{0,1\}} - P_{\{0\}} - P_{\{1\}}$. 


There is a coherence-projector $\omega_I$ for each subset of slits $I \subseteq \{1, \ldots, N\}$, defined in terms of the physical projectors:

$$\omega_I := \sum_{\tilde{I} \subseteq I} (-1)^{|\tilde{I}| + |I|} p_{\tilde{I}}.$$ 

If we demand that any state (indeed, any vector in the vector space generated by the states) in a theory can be decomposed in a form reminiscent of a rank $k$ tensor:

$$|s\rangle = \sum_{I, |I|=1}^{k} \omega_I |s_I\rangle,$$ 

then that theory has maximal order of interference $k$. This decomposition can be thought of as a generalised superposition, as it manifestly describes the coherences between different subsets of perfectly distinguishable states (the analogue of a basis in quantum theory) present in a given state.

Given this representation, we can now state the main results of [28, 30] without mention of physical principles.

**New Theorem (Extends theorem 3.0.4 in [30] and theorem 1 in [28]).** In any theory were states can be represented as a rank-$k$ tensor as in equation (3.2), for $k$ a constant, the **parity** problem for a function $f : \{1, \ldots, N\} \rightarrow \{0, 1\}$ requires a minimum of $\lceil N/k \rceil$ queries to solve, and the search problem on a database of size $N$ can be solved with $\Omega(\sqrt{N/k})$ queries.

Hence, theories where states can be represented as generalised superpositions do not provide a computational advantage over quantum theory. Although generalised superpositions may provide a constant speed-up for certain problems, this advantage does not lead to improved scaling with problem size.

### 3.3. Turing machine model

The original formulation of quantum and classical computation was in terms of (quantum and deterministic) Turing machines, rather than circuits. Can a model of computation in an arbitrary theory be formulated that utilises Turing machines? In this section, building on work in [26], we present such a generalised Turing machine model.

We start by introducing a new type of Turing machine, generalising the standard notion of a probabilistic Turing machine. In this new model, transitions can occur with quasi-probabilistic weights—rather than the standard transition probabilities occurring in probabilistic Turning machines—with the constraint that the total weight of transitions from a given state must sum to $+1$. We refer to this model as an affine Turing machine (AFTM) [26]. It was shown in [26] that the class of problems which can be efficiently solved by this model with bounded error is exactly equal to the class **AWPP**. Hence, by replacing probabilities with quasi-probabilities, one gets from the well known class **BPP** to **AWPP**, thus providing an intuitive interpretation of **AWPP**.

More formally, an AFTM is defined to be a non-deterministic Turing machine in which every transition is associated with a real-valued—not necessarily positive—weight. The weight of a given computational branch corresponds to the product of the weights of the transitions involved. It’s required that for each symbol being read by the tape head, the total weight of transitions from a given (non-halting) state must be $+1$.
Given the above, the class of languages an AffTM can efficiently decide with bounded error can now rigorously be defined. First, given an AffTM $M$ whose branches all halt in a finite number of steps, define the acceptance weight $\alpha_M(x)$ of $M$ on an input $x$ to be the total weight of the accepting paths on input $x$. An AffTM $M$ is said to be proper if $0 \leq \alpha_M(x) \leq 1$ for all inputs. It decides a language $L$ with bounded error if $\frac{1}{3} \leq \alpha_M(x) \leq 1$ for $x \in L$, and $0 \leq \alpha_M(x) \leq \frac{1}{3}$ for $x \notin L$. An AffTM is efficient if the number of transitions in any computational path on an arbitrary input $x$ is bounded from above by some polynomial in $|x|$. Given these definitions, we can now formally state the relation between AffTMs and AWPP.

**Theorem 7 (AffTM characterisation of AWPP [26]).** The class of languages decided with bounded error by some efficient proper AffTM is equal to AWPP.

This correspondence will be used to provide a Turning machine model of computation for an arbitrary theory—equivalent to the uniform circuit family model discussed in section 3.1—in terms of an AffTM satisfying one further constraint.

Fix a theory and consider a language $L$ that can be efficiently decided by it with bounded error. We saw in section 3.1 that $L \in$ AWPP, hence there exists a proper AffTM that can efficiently decide $L$ with bounded error. Write the ‘affine vector’ of this AffTM at a given time as a quasi-probability distribution overall configurations—with the term in quasi-distribution associated to a given configuration corresponding to the weight of that configuration—of the AffTM at that point in the computation. As the uniform circuit family in the theory simulates this AffTM, the real vector corresponding to the state of the theory at a given point in the circuit is an alternate representation of the affine vector at the equivalent point in the AffTM computation.

Recall from section 2 that in each theory the set of fiducial measurements can be combined into a single informationally complete measurement in such a way that the each state corresponds to the vector listing the outcome probability for each effect in this informationally complete measurement. Hence, the Euclidean norm, or 2-norm, of each state must be bounded from above by 1. Thus the two-norm of each affine vector—that is quasi-distribution over configurations—of the AffTM simulating a given circuit in the theory must also be bounded above by 1.

Hence, for a given language efficiently decided with bounded error by a specific theory $G$, there exists an efficient proper AffTM for which the Euclidean norm of each of its affine vectors is upper bounded by 1, which also decides this language with bounded error. This provides an alternate characterisation of BGP to the one provided in definition 5 which makes use of Turning machines rather than uniform circuit families.

4. Achieving the upper bound: free vs non-free theories

A natural question is whether there exits a theory such that the bound of theorem 6 is achieved. In fact, reference [26] has provided a complexity-theoretic argument that suggests this is unlikely. Despite this, by slightly modifying the definition of what constitutes a generalised probabilistic theory, one can indeed construct a theory within this altered framework—satisfying both tomographic locality and causality—in which the class of efficiently

---

13 This follows from the fact that the state with outcome probability 1 for a given effect (and hence probability 0 for every other effect due to normalisation of probabilities) maximizes [59] every other state. The two-norm of this state is clearly 1. As the two-norm is a Schur-convex function [59] the two-norm of every other state is hence bounded from above by 1.
solvable problems exactly equals $\text{AWPP}$ [26]. As we will see below, rather than assigning probabilities to any experiment composed of laboratory devices, a theory in the modified framework only assigns probabilities to certain \textit{allowed} experiments [26]. This result constitutes a converse to theorem 6 and provides an intuitive interpretation of $\text{AWPP}$, which can be thought of as the class of all problems efficiently solvable by tomographically local physical theories. The theory constructed in reference [26] has the maximum computational power consistent with tomographic locality. In a sense, one can think of it as the analogue of a PR-box—which exhibits the strongest non-local correlations consistent with the no-signalling principle—for computation.

The standard definition of a generalised theory, outlined in section 2, holds that a theory specifies a set of experimental or laboratory devices which can be composed together in sequence and parallel to form closed circuits and assigns a probability distribution over the possible outcomes of each closed circuit. Additionally, the set of devices—and device outcomes—is closed under such sequential and parallel composition. Reference [26] referred to such theories as \textit{free} generalised probabilistic theories. One can consider a modified definition of a theory, which specifies a set of devices, a set of \textit{allowed} closed circuits which can be built from those devices, and assigns a probability distribution over the outcomes of allowed closed circuits. Note that probability distributions are only assigned to the set of \textit{allowed} closed circuits specified by the theory. Such theories were referred to by reference [26] as \textit{non-free} theories. In non-free theories, states, transformations, and effect are still represented by vectors in a real vector space, matrices acting on this space, and vectors in the dual space respectively [34].

Before proceeding, it is natural to ask whether non-free theories constitute a sensible class operationally defined theories. Indeed, a standard assumption made in quantum information theory is that any circuit of unitaries can be constructed. The fact that non-free theories explicitly deny this assumption may strike some as unnatural. However, non-free theories are not as unmotivated as they may seem. All that is involved in an experiment is to start some process in motion, and at some point intervene to make an observation of some kind. Suppose that ultimately, evolution in our Universe is governed by the equations of the Standard Model. Then the Hamiltonian of the Standard Model is fixed, and one does not have the option of changing it in order to realise an arbitrary Hamiltonian evolution. The fact that certain subsystems can be identified (say, some ions in some trap, say) and their evolution controlled in an arbitrary way by placing them into a suitable environment (by applying lasers to the ions, say), is rather special, and it is as much to do with the initial condition as with the laws of the Universe. Thus the statement that an experimenter can start processes in motion, intervening at some point to make an observation, and that the process can be abstractly represented as a sequence of gates is not at odds with the assertion that there may be no physical process corresponding to an arbitrary sequence of the same gates. Hence, non-free theories are not unmotivated if one takes the viewpoint that a physical theory corresponds both to a consistent account of experimental data and to which experiments are implementable in principle.

Given the definition of non-free theories, reference [26] proved the following theorem.

\textbf{Theorem 8. } There exists a non-free theory $G$, satisfying tomographic locality and causality, such that $\text{BGP} = \text{AWPP}$.

In reference [26], it was shown that uniform poly-size circuits, in which the gates are certain affine transformations, can simulate—and be simulated by—a given AffTM. Hence, each language decided with bounded error by some efficient proper AffTM can also be decided with bounded error by a uniform circuit family built from a certain finite set of affine gates. The non-free theory of the above theorem is constructed from the collection of affine circuits.
which each AffTM, hence all of AWPP. The non-trivial part of the proof of theorem 8 is showing the non-free theory constructed in this manner satisfies tomographic locality.

5. A new conjecture

As BQP is believed to be strictly contained in AWPP [26], theorem 8 constitutes evidence against a conjecture originally made in reference [6] which posited that a quantum computer may be able to simulate computation in any generalised probabilistic theory with at most polynomial overhead. The distinction between free and non-free theories appears to capture an important aspect of the computational abilities of generalised probabilistic theories. The crucial distinction between free and non-free theories is that transformations in free theories are closed under composition, implying a bound on the set of states. Indeed, this is the intuitive rationale behind the constraint that the affine vectors of any AffTMs simulating a free theory must have two-norm bounded by 1; as the state space is bounded, the length of the affine vector cannot increase beyond a certain limit over the course of the computation. This need not be the case in non-free theories. Could a quantum computer exploit this fact and efficiently simulate computation in all tomographically local free theories? If such a conjecture was borne out, it could shed light on which physical and structural features give rise to the quantum computational speed-up. The conjunction of theorems 6 and 8 suggests the following refinement of the conjecture originally made in reference [6]. Indeed, the results outlined in section 3.2 provide weak evidence for the above conjecture. Moreover, previous results have shown that quantum theory can simulate reversible computation in Boxworld [60] and in a theory where wires in computational circuits are represented by \(d\)-dimensional Bloch balls, for \(d \neq 3\), rather than the three dimensional Bloch ball of qubits in quantum computation [61]. Both of these simulated settings satisfy tomographic locality.

New conjecture (Refinement of conjecture 2 from [6]). A quantum computer can simulate computation in any free theory satisfying \(n\)-local tomography, for \(n\) a fixed constant, with at most polynomial overhead.

6. Discussion and conclusion

This paper has reviewed and extended recent results which have explored some connections between computation and physical principles in the framework of generalised probabilistic theories [23, 25–30]. The main focus has been on understanding how simple physical principles bound the power of different computational paradigms, a first step towards understanding the source of quantum theories computational power [62]. In section 3.1, we extended a result of reference [25] by showing that in any theory satisfying \(n\)-local tomography the class of problems that can be solved efficiently is contained in AWPP—the best known bound on the power of quantum computation [46].

These results raised the question of whether quantum theory is powerful for computation in the space of all theories. Given a specific theory, under what conditions can computation in this theory be simulated by a quantum computer? Put differently, can BQP, the class of problems a quantum computer can efficiently solve, be characterised in terms of physical principles alone? Such a characterisation would deepen our understanding of quantum computation and its ultimate limitations. One can interpret the derivation of Grover’s quadratic lower bound to the search problem [28], discussed in section 3.2, as the analogue of the derivation of Tsirelson’s bound on quantum correlations from physical principles [63, 64]. A theory-independent
characterisation of BQP would then amount to the analogue of a characterisation of the entire set of quantum correlations in terms of natural principles.

While the results discussed in this paper may have deepened our understanding of quantum computers and their limits, they have not explicitly resulted in any practical applications in the same way that studying Boxworld type correlations led to the development of device-independent cryptography [65]. Can studying computation in general theories different from quantum theory result in practical applications? One potential avenue for this is blind and verified delegated computation [66].

Consider the situation where a computationally bounded client wants to delegate her computation to a server with access to a full quantum computer. The protocol for blind computation provided in [66] ensures that the client can have a server carry out a quantum computation for her such that the client’s inputs, outputs, and computation remain perfectly private. Hence, a malicious server cannot learn any of the client’s information, and all the client needs to be able to do is prepare single qubit states and send them to the server. Moreover, the security of this protocol has been shown to follow from the no-signalling principle [67]. However, while the server cannot learn the client’s computation, they can still tamper with it by deviating from the client’s instructions. Using a scheme introduced by reference [66], the client can detect any deviations made by the server with probability exponentially close to one.

However, the correctness of this verification protocol rests on the assumption that the server can only deviate from the specified instructions by using quantum dynamics. The client may not be able to detect deviations which use post-quantum dynamics. Hence, as was the case for quantum key distribution before the work of Barrett, Hardy, and Kent [65], it is not clear if delegated quantum computation is secure against post-quantum attacks. This raises the question of whether the correctness of this verification protocol can be established directly from physical principles. Indeed, a derivation of BQP from physical principles would lay the groundwork for a proof of the correctness of delegated computation solely from first principles.

One can already provide some initial analysis using known results in the literature. For instance, as the protocol of [66] requires the client to prepare and communicate single qubit states, one can wonder what additional requirements are needed to ensure the client can detect deviations when the server is not assumed to employ quantum dynamics a priori. In order for it to make good operational sense, we assume the theory used by the server must belong to the framework discussed here. Reference [43] has shown that quantum theory is the unique generalised probabilistic theory theory where the local systems are qubits, global systems obey tomographic locality, and where there exists at least one continuous interaction between systems. Hence, if one assumes such structure, then the standard correctness of delegated computation discussed above applies. We have replaced the requirement that the server is bound by quantum theory by the assumption of tomographic locality, and the existence of a reversible, continuous interaction between subsystems—as long as local systems are qubits. In fact, one can go further. Recent work in Reference [61] has shown that if local systems are $d$-dimensional Bloch balls (a qubit, for instance, lives in a three dimensional Bloch ball), tomographic locality and causality are satisfied, and the global transformations form a closed, connected matrix.

---

14 It should be pointed out that this problem becomes non-trivial only when the communication between verifier and prover involves communicating quantum states, as in the protocol by reference [66] discussed here. When the communication between a single prover and verifier is classical, then all standard results in interactive proof systems hold, since they apply to arbitrary provers.

15 Indeed, the server could simply refuse to perform the client’s computation. Such a refusal is immediately obvious to the client. As nothing can be done to force the server to perform the computation, we only consider situations in which the server does carry out a computation. It is then up to the client to verify if the performed computation is the one specified.
then the only theory with non-trivial interactions is when $d = 3$. That is, when local systems are qubits. All other cases only have local transformations, and hence can be simulated on a classical computer.

In the above two instances we have been able to replace the requirement that the server be bound by quantum theory by other—and in the second case, weaker—assumptions. If one has a full derivation of BQP from physical principles, then one could perhaps replace the requirement that the server be bound by quantum theory in the correctness proof by the requirement the server satisfy these physical principles. This would place the correctness and security of delegated computation on par with physical principles, in the same way that Barrett, Hardy, and Kent [65] placed security of quantum key distribution on par with the no-signalling principle.

Acknowledgments

The author thanks Jonathan Barrett and John Selby for useful discussions. This project was funded by the EPSRC through the National Quantum Technology Hub in Networked Quantum Information Technologies and the UCL doctoral prize fellowship (Project Number: 534936).

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Ciarán M Gilligan-Lee https://orcid.org/0000-0002-7038-3954

References

[1] Peres A 1995 Quantum Theory: Concepts and Methods (Dordrecht: Kluwer)
[2] Hardy L 2005 Probability theories with dynamic causal structure: a new framework for quantum gravity (arXiv:gr-qc/0509120)
[3] Hardy L 2011 Reformulating and reconstructing quantum theory (arXiv:1104.2066)
[4] Chiribella G, D’Ariano G M and Perinotti P 2010 Probabilistic theories with purification Phys. Rev. A 81 062348
[5] Chiribella G, D’Ariano G M and Perinotti P 2011 Informational derivation of quantum theory Phys. Rev. A 84 012311
[6] Barrett J 2007 Information processing in generalized probabilistic theories Phys. Rev. A 75 032304
[7] Barnum H, Lee C, Scandolo C and Selby J 2017 Ruling out higher-order interference from purity principles Entropy 19 253
[8] Lee C M and Selby J H 2017 A no-go theorem for theories that decohere to quantum mechanics (arXiv:1701.07449)
[9] Lee C 2018 Beyond quantum New Sci. 238 28–9
[10] Galley T D and Masanes L 2016 Classification of all alternatives to the born rule in terms of informational properties (arXiv:1610.04859)
[11] D’Ariano G M, Manessi F, Perinotti P and Tosini A 2014 Fermionic computation is non-local tomographic and violates monogamy of entanglement Europhys. Lett. 107 20009
[12] D’Ariano G M, Manessi F, Perinotti P and Tosini A 2014 The Feynman problem and fermionic entanglement: Fermionic theory versus qubit theory Int. J. Mod. Phys. A 29 1430025
[13] Wootters W K 1990 Local accessibility of quantum states Complexity, Entropy and the Physics of Information ed W H Zurek (Addison: Westview Press) pp 39–46
[14] D’Ariano G, Manesso F and Perinotti P 2014 Determinism without causality Phys. Scr. **T163** 014013
[15] Massar S, Pironio S and Pitalúa-García D 2015 Hyperdense coding and superadditivity of classical capacities in hypersphere theories *New J. Phys.* **17** 113002
[16] Janotta P, Gogolin C, Barrett J and Brunner N 2011 Limits on nonlocal correlations from the structure of the local state space *New J. Phys.* **13** 063024
[17] Spekkens R W 2007 Evidence for the epistemic view of quantum states: a toy theory *Phys. Rev.* A **75** 032110
[18] Janotta P and Lal R 2013 Generalized probabilistic theories without the no-restriction hypothesis *Phys. Rev.* A **87** 052131
[19] Popescu S and Rohrlich D 1998 Causality and nonlocality as axioms for quantum mechanics *Causality and Locality in Modern Physics* (Berlin: Springer) pp 383–9
[20] Popescu S 2014 Nonlocality beyond quantum mechanics *Nat. Phys.* **10** 264–70
[21] Cavalcanti P J, Selby J H, Sikora J, Galley T D and Sainz A B 2021 Witworld: a generalised probabilistic theory featuring post-quantum steering (arXiv:2102.06581)
[22] Lee C M and Hoban M J 2016 The information content of systems in general physical theories (arXiv:1606.06801)
[23] Lee C M and Selby J H 2016 Generalised phase kick-back: the structure of computational algorithms from physical principles *New J. Phys.* **18** 033023
[24] Barnum H et al 2014 Higher-order interference and single-system postulates characterizing quantum theory *New J. Phys.* **16** 123029
[25] Lee C M and Barrett J 2015 Computation in generalised probabilistic theories *New J. Phys.* **17** 083001
[26] Barrett J, de Beaudrap N, Hoban M J and Lee C M 2019 The computational landscape of general physical theories *npj Quantum Inf.* **5** 1–10
[27] Lee C M and Hoban M J 2016 Bounds on the power of proofs and advice in general physical theories *Proc. R. Soc., A* **472** 20160076
[28] Lee C M and Selby J H 2016 Deriving Grover’s lower bound from simple physical principles *New J. Phys.* **18** 093047
[29] Lee C M and Selby J H 2017 Higher-order interference in extensions of quantum theory *Found. Phys.* **47** 89–112
[30] Lee C M, Selby J H and Barnum H 2017 Oracles and query lower bounds in generalised probabilistic theories (arXiv:1704.05043)
[31] Abramsky S and Coecke B 2004 A categorical semantics of quantum protocols 2004 *Proc. 19th Annual IEEE Symp. on Logic in Computer Science* (Piscataway, NJ: IEEE) pp 415–25
[32] Coecke B and Kissinger A 2016 *Picturing Quantum Processes: A First Course in Quantum Theory and Diagrammatic Reasoning* (Cambridge: Cambridge University Press)
[33] Selby J H and Lee C M 2020 Compositional resource theories of coherence *Quantum* **4** 319
[34] Lee C M 2016 Bounds on computation from physical principles *DPhil Thesis* University of Oxford
[35] Prugovecki E 1977 Information-theoretical aspects of quantum measurement *Int. J. Theor. Phys.* **16** 321–31
[36] Busch P and Lahti P J 1989 The determination of the past and the future of a physical system in quantum mechanics *Found. Phys.* **19** 633–78
[37] Masanes L, Müller M P, Augusiak R and Pérez-García D 2013 Existence of an information unit as a postulate of quantum theory *Proc. Natl Acad. Sci.* **110** 16373–7
[38] Flammia S T, Silberfarb A and Caves C M 2005 Minimal informationally complete measurements for pure states *Found. Phys.* **35** 1985–2006
[39] Coecke B 2014 Terminality implies non-signalling (arXiv:1405.3681)
[40] Hardy L and Wootters W K 2012 Limited holism and real-vector-space quantum theory *Found. Phys.* **42** 454–73
[41] Morimae T and Nishimura H 2015 Power of quantum computing with restricted postselections (arXiv:1502.00067)
[42] Morimae T, Nishimura H and Gall F L 2016 Modified group non-membership is in awpp (arXiv:1602.06073)
[43] de Beaudrap N 2014 On computation with probabilities’ modulo k (arXiv:1405.7381)
[44] Bremner M J, Jozsa R and Shepherd D J 2010 Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy *Proc. R. Soc. A* **467** 459–72
[45] Chaitin G J 1975 A theory of program size formally identical to information theory *J. ACM* **22** 329–40
[46] Fortnow L and Rogers J 1998 Complexity limitations on quantum computation 1998 Proc. 13th Annual IEEE Conf. Computational Complexity (Piscataway, NJ: IEEE) pp 202–9
[47] Bennett C H, Bernstein E, Brassard G and Vazirani U 1997 Strengths and weaknesses of quantum computing SIAM J. Comput. 26 1510–23
[48] Grover L K 1997 Quantum mechanics helps in searching for a needle in a haystack Phys. Rev. Lett. 79 325
[49] Nielsen M A and Chuang I L 2010 Quantum Computation and Quantum Information (Cambridge: Cambridge University Press)
[50] Meyer D A and Pommersheim J 2011 On the uselessness of quantum queries Theor. Comput. Sci. 412 7068–74
[51] Sorkin R D 1994 Quantum mechanics as quantum measure theory Mod. Phys. Lett. A 09 3119–27
[52] Barnum H, Barrett J, Krumm M and Müller M P 2015 Entropy, majorization and thermodynamics in general probabilistic theories (arXiv:1508.03107)
[53] Niestegge G 2016 Quantum teleportation and Grover’s algorithm without the wavefunction (arXiv:1611.02926)
[54] Dakić B, Paterek T and Brukner Č 2014 Density cubes and higher-order interference theories New J. Phys. 16 023028
[55] Gogioso S and Scandolo C M 2018 Density hypercubes, higher order interference and hyper-decoherence: a categorical approach Int. Symp. on Quantum Interaction (Berlin: Springer) pp 141–60
[56] Hefford J and Gogioso S 2020 Hyper-decoherence in density hypercubes (arXiv:2003.08318)
[57] Sorkin R D 1995 Quantum measure theory and its interpretation (arXiv:gr-qc/9507057)
[58] Sinha A, Vijay A H and Sinha U 2015 On the superposition principle in interference experiments Sci. Rep. 5 10304
[59] Marshall A W, Ollin I and Arnold B C 1979 Inequalities: Theory of Majorization and its Applications vol 143 (Berlin: Springer)
[60] Gross D, Müller M, Colbeck R and Dahlsten O C 2010 All reversible dynamics in maximally nonlocal theories are trivial Phys. Rev. Lett. 104 080402
[61] Krumm M and Müller M P 2019 Quantum computation is the unique reversible circuit model for which bits are balls npj Quantum Inf. 5 1–8
[62] Lee C 2019 Heart of quantum New Sci. 241 38–41
[63] Pawlowski M, Paterek T, Kaszlikowski D, Scarani V, Winter A and Zukowski M 2009 Information causality as a physical principle (arXiv:0905.2292)
[64] Fritz T, Sainz A, Augusiak R, Brask J B, Chaves R, Leverrier A and Acín A 2013 Local orthogonality as a multipartite principle for quantum correlations Nat. Commun. 4 2263
[65] Barrett J, Hardy L and Kent A 2005 No signaling and quantum key distribution Phys. Rev. Lett. 95 010503
[66] Broadbent A, Fitzsimons J and Kashefi E 2009 Universal blind quantum computation 2009 FOCS'09 50th Annual IEEE Symp. on Foundations of Computer Science (Piscataway, NJ: IEEE) pp 517–26
[67] Morimae T and Fujii K 2013 Blind quantum computation protocol in which alice only makes measurements Phys. Rev. A 87 050301