General Conditions for Universality of Quantum Hamiltonians

Tamara Kohler

QMath summer school, August 2024

1 Recap

Recall from Toby's lectures, we define approximate simulation as:

Definition 1.1 (Approximate simulation, [Cubitt et al., 2018, Def. 23])**.** *Let $\Delta, \eta, \epsilon > 0$. A Hamiltonian H' is a (Δ, η, ϵ)-simulation of the Hamiltonian H if there exists a local encoding $\mathcal{E}(M) = V(M \otimes P + \overline{M} \otimes Q)V^\dagger$ such that*

i. There exists an encoding $\tilde{\mathcal{E}}(M) = \tilde{V}(M \otimes P + \overline{M} \otimes Q)\tilde{V}^\dagger$ into the subspace $S_{\tilde{\mathcal{E}}}$ such that $S_{\tilde{\mathcal{E}}} = S_{\leq \Delta(H')}$ and $\|\tilde{V} - V\| \leq \eta$; and

ii. $\|H'_{\leq \Delta} - \tilde{\mathcal{E}}(H)\| \leq \epsilon$.

where an encoding is a map of the form:

$$\mathcal{E}(A) = V\left(A \otimes P + \overline{A} \otimes Q\right)V^\dagger, \tag{1}$$

and a *local* encoding is an encoding which maps local observables to local observables, defined as follows.

Definition 1.2 (Local encoding (Definition 13 from [Cubitt et al., 2018]))**.** *Let*

$$\mathcal{E} : \mathcal{B}\left(\otimes_{j=1}^n \mathcal{H}_j\right) \to \mathcal{B}\left(\otimes_{j=1}^n \mathcal{H}'_j\right)$$

be an encoding. We say that the encoding is local if for any operator $A_j \in \mathrm{Herm}(\mathcal{H}_j)$ there exists $A'_j \in \mathrm{Herm}(\mathcal{H}'_j)$ such that:

$$\mathcal{E}(A_j \otimes \mathbb{1}) = (A'_j \otimes \mathbb{1})\mathcal{E}(\mathbb{1}).$$

We say a family of Hamiltonians $\mathcal{M}$ is a universal simulator if any finite dimensional Hamiltonian can be simulated by a Hamiltonian from $\mathcal{M}$.

As Toby will prove in his lectures, for 2-qubit interactions, the set of families of Hamiltonians which are universal coincides with the set of families for which the ground state energy problem is QMA-complete. Is this just a coincidence?

2 Classification of universality for classical Hamiltonians

In the classical case it is known that the set of universal families of Hamiltonians can be completely characterised by complexity classes [De las Cuevas and Cubitt, 2016].

More precisely - a classical spin model is universal (in that it can simulate all classical spin physics) iff it is closed and its ground state energy problem admits a faithful, polynomial time reduction from SAT.

Where we say a family of Hamiltonians $\mathcal{M}$ is closed if given $H_A^{(1)} \in \mathcal{M}$ and $H_B^{(2)} \in \mathcal{M}$ there exists $H \in \mathcal{M}$ that simulates $H_A^{(1)} + H_B^{(2)}$.

And (informally) the ground state energy problem of a family of Hamiltonians admits a faithful, polynomial time reduction from SAT if given any, arbitrary instance of SAT there is a Hamiltonian from the family that has a low energy ground state iff the SAT instance is satisfiable and low energy ground states of the Hamiltonian are in 1-to-1 correspondence with the satisfying SAT assignments.

If the ground state energy problem of a family of Hamiltonians admits a faithful polynomial time reduction from SAT then the ground state energy problem of the family of Hamiltonians is NP-complete. So there are some similarities between the classification of universal 2-qubit Hamiltonians and universal classical Hamiltonians. But the proof techniques are *completely* different. In the classical case the proof uses the fact that the SAT reduction allows us to encode classical computation into the structure of the ground state. In the quantum case the proof uses a sequence of perturbative reductions. Can we use the intuition from the classical case to obtain a complexity theoretic classification of universal quantum simulators?

3 Translationally invariant universal Hamiltonians in 1D

3.1 Circuit-to-Hamiltonian mappings

There is a way to encode quantum computation into the ground state of a local quantum Hamiltonian. This technique was first proposed by Feynman in 1986, and is the foundation for many prominent results in Hamiltonian complexity theory, such as QMA-hardness of the local Hamiltonian problem [Feynman, 1985, Kitaev et al., 2002].

These are often called "circuit-to-Hamiltonian mappings", though the mappings may involve other models of quantum computation than the circuit model. These Hamiltonians are typically constructed in such a way that their ground states are "computational history states". A very general definition of history states was given in [González-Guillén and Cubitt, 2018]; we will only require the simpler "standard" history states here:

Definition 3.1 (Computational history state)**.** *A computational history state* $|\Phi\rangle_{QC} \in \mathcal{H}_Q \otimes \mathcal{H}_C$ *is a state of the form*

$$|\Phi\rangle_{QC} = \frac{1}{\sqrt{T}} \sum_{t=1}^{T} |\psi_t\rangle |t\rangle ,$$

where $\{|t\rangle\}$ *is an orthonormal basis for* $\mathcal{H}_C$ *and* $|\psi_t\rangle = \Pi_{i=1}^{t} U_i |\psi_0\rangle$ *for some initial state* $|\psi_0\rangle \in \mathcal{H}_Q$ *and set of unitaries* $U_i \in \mathcal{B}(\mathcal{H}_Q)$*.*

$\mathcal{H}_C$ *is called the clock register and* $\mathcal{H}_Q$ *is called the computational register. If* U_t *is the unitary transformation corresponding to the* t^{th} *step of a quantum computation, then* $|\psi_t\rangle$ *is the state of the computation after* t *steps. We say that the history state* $|\Phi\rangle_{QC}$ *encodes the evolution of the quantum computation.*

Note that U_t need not necessarily be a gate in the quantum circuit model. It could also e.g. be one time-step of a quantum Turing machine, or even a time-step in some more exotic model of quantum computation [Bausch et al., 2017], or an isometry [Usher et al., 2017]. In the particular constructions we make use of in this section, U_t will be a time-step of a quantum Turing machine.

What Hamiltonian has this as its ground state?

$$H = H_{\text{in}} + H_{\text{clock}} + H_{\text{prop}} \tag{2}$$

where H_{in} enforces that when the clock is in its start state, the computational register is initialised in the correct state, H_{clock} ensures that the clock part of the Hilbert space is always encoding a valid time state and H_{prop} enforces that as the clock ticks the computational qubits evolve under the desired circuit.

The exact form of the Hamiltonian depends on what type of quantum computation is being encoded. In our case the Hamiltonian terms are quite complicated - we won't go through them in detail (see [Kohler et al., 2020] for info) - but we'll go through a simpler example.

If the model of computation being encoded is the circuit model and the clock is a unary clock encoded into qubits then the Hilbert space decomposes as $\mathcal{H} = \mathcal{H}_Q \otimes \mathcal{H}_C = (\mathbb{C}^2)^n \otimes (\mathbb{C}^2)^{\otimes T+1}$ where T is the number of steps in the computation.

The start time of the computation is given by the clock state $|0\rangle^{\otimes T+1}$. The second time step of the computation is given by the clock state $|1\rangle |0\rangle^{\otimes T}$. ... The penultimate time step of the computation has clock state $|1\rangle^{\otimes T} |0\rangle$. The final time step of the computation has clock state $|1\rangle^{\otimes T+1}$.

The initialisation term is simply $H_{\text{in}} = \sum_{j \in A} |0\rangle\langle 0|_j \otimes |1\rangle\langle 1|_0^{(c)}$ where A is the set of computational qubits that are initially initialised to the $|0\rangle$ state (some other set of computational qubits – the qubits that form the witness for a verification circuit – will not be initialised, these computational qubits can be in any state at the start of the circuit).

The clock term is:

$$H_{\text{clock}} = \sum_{t=0}^{T} |01\rangle\langle 01|^{(c)}_{t,t+1}$$

This term gives a penalty to any clock term that isn't of the form $|1\rangle^n |0\rangle^{T+1-n}$ – i.e. any clock term that isn't a valid state of the unary clock.

The propagation term is:

$$H_{\text{prop}} = \frac{1}{2} \sum_{t=0}^{T} H_t$$

where:

$$H_t = \mathbb{1} \otimes \Big(|10\rangle\langle 10|^{(c)}_{t,t+1} + |10\rangle\langle 10|^{(c)}_{t+1,t+2} \Big) - U_{t+1} \, |110\rangle\langle 100|^{(c)}_{t-1,t,t+1} - U^{\dagger}_{t+1} \, |100\rangle\langle 110|^{(c)}_{t-1,t,t+1}$$

3.2 Quantum Turing machines

The type of computation we will be encoding into the ground state of our Hamiltonian is a quantum Turing machine. A Turing machine is a mathematical model of computation describing an abstract machine that manipulates symbols on a tape according to a set of rules.

Definition 3.2 (Turing machine)**.** *A Turing machine is a tuple* (Q, Σ, δ)*. Where:*

- $Q = \{q_0, q_1, \ldots, q_f\}$ *is some finite set of states*
- $\Sigma = \{\sigma_0, \sigma_1, \ldots, \perp\}$ *is a finite set of symbols, with a distinguished blank symbol* $\perp$
- $\delta : Q \times \Sigma \to Q \times \Sigma \times \{L, R\}$ *is a partial function from state, symbol pairs to state, symbol direction triples*

How does a Turing machine work?

- Start:
  - head over tape cell 0
  - input: finite string of non-blank symbols written on tape from cell 0 onwards
  - rest of tape filled with blank symbols
  - Turing machine in state q_0
- Each time step

Track	Purpose
1	Input track, contains input state $\|\psi\rangle \in \mathbb{C}^2$ followed by string of $\|0\rangle$s
2	Turing machine work tape
3	Tape head and state
4	Clock tracks

Table 1: Local Hilbert space decomposition

  - tape symbol σ under head, Turing machine currently in internal state q
  - partial function is evaluated $\delta(\sigma, q) = (\sigma', q', D)$ for $D \in \{L, R\}$
  - write σ' on cell under head, update internal state to q', move head one step in direction D

- Finish
  - if Turing machine enters state q_f halt
  - output: whatever is left written on tape

A quantum Turing machine is a generalisation of a classical Turing machine, where now the set of states Q is a Hilbert space $\mathcal{H}_Q$, the set of symbols Σ is replaced by a Hilbert space $\mathcal{H}_\Sigma$, and the partial function δ is replaced by a unitary U_δ. For a more rigorous definition see [Bernstein and Vazirani, 1997].

How do we map a quantum Turing machine to a Hamiltonian? We won't map into a system of qubits, instead we'll map to a system of high dimensional qudits. The state space of the each qudit can be split into four tracks with different purposes. It is shown in [Gottesman and Irani, 2009b] that there exists a translationally invariant, nearest neighbour Hamiltonian acting on a chain of qudits (with state space decomposition as shown in Table 1). For details of the construction refer to [Gottesman and Irani, 2009b] – intuition is that in a Turing machine computations are translationally invariant – the same unitary is applied at every point (very different to the circuit model).

All the 'circuit-to-Hamiltonian' mappings we make use of in this work are what are known as "Standard form Hamiltonians". Where "Standard form Hamiltonians" are a certain class of circuit-to-Hamiltonian constructions, defined in [Watson, 2019]. We refer interested readers to [Watson, 2019] for the full definition - and simply note that it encompasses the Turing-machine based mappings which we make use of in this work [Cubitt et al., 2015, Gottesman and Irani, 2009b]. In [Watson, 2019], the following result was shown, which we will make use of in our proofs:

Lemma 3.3 (Standard form ground states; restatement of [Watson, 2019, Lem. 5.8, Lem. 5.10])**.** *Let H_{SF} be a Standard Form Hamiltonian encoding a computation U, which takes (classical) inputs from a Hilbert space $\mathcal{S}$, and*

which sets an output flag with certainty if it is given an invalid input. For $|\psi_\mu\rangle \in \mathcal{S}$ *and* $\Pi_{t=1}^{T} U_t = U$ *we define*

$$|\Phi(U,\psi_\mu)\rangle \coloneqq \frac{1}{\sqrt{T}} \sum_{t=1}^{T} U_t \ldots U_1 |\psi_\mu\rangle |t\rangle .$$

Then $\mathcal{L} = \mathrm{span}\{|\Phi(U,\psi_\mu)\rangle\}_{\mu=1}^{d^n}$ *defines the kernel of* H_{SF}*, i.e.* $H_{\mathrm{SF}}|_{\mathcal{L}} = 0$*. The smallest non-zero eigenvalue of* H_{SF} *scales as* $1 - \cos \pi/2T$*.*

3.3 High-level outline of the construction

In both the constructions we are going to discuss today a description of the Hamiltonian to be simulated (the "target" Hamiltonian, H_{target}) is encoded in the binary expansion of some natural number, $x \in \mathbb{N}$. Before looking at how we encode x into the parameter of the simulator Hamiltonian, we first look at how we encode the target Hamiltonian into the binary expansion of x

We will assume that $H_{\mathrm{target}} = \sum_i h'_i$ is a k-local Hamiltonian, acting on n spins of local dimension d. We emphasize that k can be taken to be n, i.e. the system size—and therefore we can simulate *any* Hamiltonian, not just local ones. However we keep track of the locality parameter k as it is relevant when deriving the overhead of our simulations.

The binary expansion $B(x)$ of x has the following form:

$$B(x) \coloneqq \gamma'(n)\cdot\gamma'(k)\cdot\gamma'(m)\cdot\left[\gamma'(i)^{\cdot k}\cdot\left(\gamma'(a_j)\cdot\gamma'(b_j)\right)^{4^k}\right]^{\cdot m}\cdot\gamma'(\Xi)\cdot\gamma'(L). \quad (3)$$

Here $\gamma'(n)$ denotes n in Elias-γ' coding, and $\cdot$ denotes concatenation of bit strings.

where:

- $\gamma'(n)$ denotes n in Elias-γ' coding, and $\cdot$ denotes concatenation of bit strings
- n denotes the total number of spins H_{target} acts on
- k is the locality of the Hamiltonian
- m is the number of terms
- i labels which spins a particular term acts on
- a, b specify the matrix values (in the form $m = a - b\sqrt{2}$)
- Ξ and L are parameters controlling the error of the simulation - more on these later

Track	Purpose
1	Input track, contains input state $\|\psi\rangle \in \mathbb{C}^2$ followed by string of $\|0\rangle$s
2	Turing machine work tape (shared by M_1 and M_{PE})
3	Tape head and state for M_1
4	Tape head and state for M_{PE}
5	Clock tracks

Table 2: Local Hilbert space decomposition for the construction.

Specifying H_{target} to accuracy δ requires each such matrix entry to be specified to accuracy $\delta/(md^{2k})$. Therefore the length of the description of H_{target} is

$$md^{2k} \log\left(\|H_{\text{target}}\| md^{2k}/\delta\right) = \text{poly}\left(n, d^k, \log(\|H\|/\delta)\right) \tag{4}$$

The natural number x is then itself encoded in some parameter of the universal Hamiltonian. We will cover two methods of doing this:

- $x = N$ where N is the number spins in the chain
- $x = \phi$ where the simulator Hamiltonian is given by $H_{\text{sim}} = \sum_i h_i$ and $h_i = A + \left(e^{i\pi\phi}B + e^{i\pi 2^{-|\phi|}}C + h.c.\right)$ where A is a fixed Hermitian matrix and B, C are fixed non-Hermitian matrices (a detailed construction of the terms in the Hamiltonian can be found in [Cubitt et al., 2015, Section 4]).

So now we have a method of encoding H_{target} into a natural number, and two methods for encoding this natural number into some parameter of a simulator Hamiltonian H_{sim}. How do we go from here to constructing an H_{sim} that simulates H_{target}? As you might have guessed, the construction relies heavily on history state Hamiltonians.

We are going to construct a Hamiltonian that has as its ground state computational history states which encode two QTMs (M_1 and M_{PE}) which share a work tape. The local Hilbert space decomposition of the system that encodes this Hamiltonian is given in Table 2 The two computations are 'dovetailed' together - the computation M_1 occurs first, and the result of this computation is used as input for M_{PE}. The first QTM, M_1, extracts the binary expansion of x from the parameter of the Hamiltonian (if $x = N$ then M_1 is a binary counter Turing machine that counts the number of spins in the chain, if $x = \phi$ then M_1 is a phase estimation Turing machine that carries out phase estimation on the terms in the Hamiltonian). At the end of M_1's computation, the binary expansion of x is written on the work tape which M_1 shares with M_{PE}.

The second QTM, M_{PE} reads in x, which contains a description of H_{target}, from the work tape which it shares with M_1. It also reads in an input state $|\psi\rangle$

- this is unconstrained by the computation (it can be thought of as carrying out the same role as a witness in a QMA verification circuit). It then carries out phase estimation on $|\psi\rangle$ with respect to the unitary generated by H_{target}. The parameter Ξ in the description of H_{target} controls the precision to which this phase estimation is carried out.

The Hamiltonian which encodes M_1 and M_{PE} has a zero-energy degenerate ground space, spanned by history states with all possible input states $|\psi\rangle$. In order to recreate the spectrum of H_{target} we need to break this degeneracy. We achieve this by adding one body projectors to the universal Hamiltonian which give the correct energy to the output of M_{PE} to reconstruct the spectrum of H_{target}.

With this construction the energy levels of the universal Hamiltonian recreate the energy levels of H_{target}. To ensure that the eigenstates are also correctly simulated, before M_1 carries out its computation, it 'idles' in its initial state for some time L (which is the final parameter in the description of H_{target}). Another way to think about why this idling is necessary is to note that history states are highly entangled – there is no *local* encoding that approximately maps from the state of the physical spins to the history state subspace because of this entanglement. But by idling before carrying out the computation we can significantly reduce the entanglement in the ground state of the simulator Hamiltonian in such a way that we can show that there does exist a local encoding which approximately maps from the state of the physical spins to the low energy subspace of the simulator system. By choosing L large enough, we show that this construction can approximately simulate any target Hamiltonian.

3.4 Translationally-invariant universal models in 1D

We can now prove our result more formally.

We will require a digital quantum simulation algorithm, summarized in the following lemm:

Lemma 3.4 (Implementing a Local Hamiltonian Unitary)**.** *For a k-local Hamiltonian $H = \sum_{i=1}^{m} h_i$ on an n-partite Hilbert space of local dimension d, and where $m = \text{poly}\, n$, there exists a QTM that implements a unitary $\tilde{U}$ such that*

$$\tilde{U} = \mathrm{e}^{\mathrm{i}Ht} + \mathrm{O}(\epsilon),$$

and which requires time $\text{poly}(1/\epsilon, d^k, \|H\|t, n)$.

Proof. Follows directly from [Lloyd, 1996, Berry et al., 2005]. □

The polynomial time bound in Lemma 3.4 suffices for our purposes; a tighter (and more complicated) bound, also for the more general case of sparse Hamiltonians, can be found in [Berry et al., 2015].

We can now start our main analysis by proving that "dovetailing" quantum computations—rigorously defined and constructed in [Cubitt et al., 2015, Lem. 22]—can be used to construct universal simulators.

Lemma 3.5 (Dovetailing for simulation)**.** *Let M_1 be a QTM which writes out the binary expansion of some $x \in \mathbb{N}$ on its work tape. Assume there exists a standard form Hamiltonian which encodes the Turing machine M_1. Then there also exists a standard form Hamiltonian $H_{\mathrm{SF}}(x)$, which encodes the computation M_1 dovetailed with a QTM M_{PE}, such that the family of Hamiltonians*

$$H_{\mathrm{univ}}(x) = \Delta H_{\mathrm{SF}}(x) + T \sum_{i=0}^{N-1} \left(\sqrt{2}\Pi_\alpha - \Pi_\beta \right) \tag{5}$$

can simulate any quantum Hamiltonian. Here Δ and T are parameters of the model, and Π_α and Π_β are one-body projectors,

Proof of Lemma 3.5. To prove this we show that the $H_{\mathrm{univ}}(x)$ can satisfy the definition to be an approximate simulation of an arbitrary "target Hamiltonian" H_{target}, to any desired accuracy. We break up the proof into multiple parts. First we construct a history state Hamiltonian $H_{\mathrm{SF}}(x)$, which encodes two Turing machine computations: M_1 which extracts a description of H_{target} from a parameter of H_{SF}, and M_{PE} which carries out phase estimation on the unitary generated by H_{target}. Then we define the one-body projectors Π_α and Π_β which break up the ground space degeneracy of H_{SF}, and inflict just the right amount of penalty to approximately reconstruct the spectrum of H_{target} in its entirety.

Construction of $\mathbf{H_{SF}}$. H_{SF} is a standard form history state Hamiltonian with a ground space laid out in Lemma 3.3. Recall that the local states of the spins on which H_{SF} acts are divided into multiple "tracks".

The QTM M_{PE} reads in the description of H_{target}—provided as integer $x \in \mathbb{N}$ output by the Turing machine M_1 whose worktape it shares. M_{PE} further reads in the unconstrained input state $|\psi\rangle$. But instead of proceeding immediately, M_{PE} idles for L time-steps (where L is specified in the input string x), before proceeding to carry out the quantum phase estimation algorithm.

The quantum phase estimation algorithm is carried out with respect to the unitary $U = \mathrm{e}^{\mathrm{i}H_{\mathrm{target}}\tau}$ for some τ such that $\|H_{\mathrm{target}}\tau\| < 2\pi$. It takes as input an eigenvector $|u\rangle$ of U, and calculates the eigenphase a_u and b_u such that the eigenphase $\phi_u = \sqrt{2}a_u - b_u$. The output of M_{PE} is then the pair of integers (a_u, b_u) specified in binary on an output track. To calculate λ_u—the eigenvalue of H_{target}—to accuracy ϵ requires determining ϕ_u to accuracy $\mathrm{O}(\epsilon/\|H_{\mathrm{target}}\|)$ which takes $\mathrm{O}(\|H_{\mathrm{target}}\|/\epsilon)$ uses of $U = \mathrm{e}^{\mathrm{i}H_{\mathrm{target}}\tau}$. The unitary U must thus be implemented to accuracy $\mathrm{O}(\epsilon/\|H_{\mathrm{target}}\|)$, which is done using

Lemma 3.4; the latter introduces an overhead $\mathrm{poly}(n, d^k, \|H_{\text{target}}\|, \tau, 1/\epsilon)$ in the system size n, local dimension d, locality k, and target accuracy ϵ. The error overhead of size poly $1/\epsilon$ due to the digital simulation of the unitary is thus polynomial in the precision, as are the $\propto 1/\epsilon$ repetitions required for the QPE algorithm. The whole procedure takes time

$$T_{\text{PE}} \coloneqq \mathrm{poly}(d^k, \|H_{\text{target}}\|/\epsilon, n). \tag{6}$$

In our construction the input to M_{PE} is not restricted to be an eigenvector of $|u\rangle$, but it can always be decomposed as $|\psi\rangle = \sum_u m_u |u\rangle$. By linearity, for input $|\psi\rangle = \sum_u m_u |u\rangle$ the output of M_{PE} will be a superposition in which the output (a_u, b_u) occurs with amplitude m_u.

After M_{PE} has finished its computation, its head returns to the end of the chain. A dovetailed counter then decrements $a_u, a_u - 1, \ldots, 0$ and $b_u, b_u - 1, \ldots, 0$.[1] For each timestep in the counter $a_u, a_u - 1, \ldots, 0$ the Turing machine head changes one spin to a special flag state $|\Omega_a\rangle$ which does not appear anywhere else in the computation. While for each timestep in the counter $b_u, b_u - 1, \ldots, 0$ the Turing machine head changes one spin to a different flag state $|\Omega_b\rangle$. (See e.g. [Bausch et al., 2018, Lem. 16]) for a construction of a Turing machine with these properties.)

By Lemma 3.3, the ground space $\mathcal{L}$ of H_{SF} is spanned by computational history states as given in Definition 3.1, and is degenerate since any input state $|\psi\rangle$ yields a valid computation. Therefore:

$$\ker(H_{\text{SF}}) = \mathcal{L} = \mathrm{span}_{|\psi\rangle}\left(\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left|\psi^{(t)}\right\rangle|t\rangle\right) \tag{7}$$

where $\left|\psi^{(t)}\right\rangle$ denotes the state of the system at time step t if the input state was $|\psi\rangle$.

A Local Encoding. In order to prove that $H_{\text{univ}}(N)$ can simulate all quantum Hamiltonians, we need to demonstrate that there exists a local encoding $\mathcal{E}(M)$ such that the conditions of Definition 1.1 are satisfied. To this end, let

$$|\Phi_{\text{idling}}(\psi)\rangle \coloneqq \frac{1}{\sqrt{L'}}\sum_{t=1}^{L'}\left|\psi^{(t)}\right\rangle|t\rangle$$

where $L' = T_1 + L$, and where T_1 is the number of time steps in the M_1 computation. This is the history state up until the point that M_{PE} begins its computation (i.e. the point at which the 'idling to enhance coherence' ends). So, throughout the computation encoded by this computation the

[1] For general input state $|\psi\rangle = \sum_u m_u |u\rangle$ there will be a superposition where the counter $a_u, a_u - 1, \ldots, 0$ and $b_u, b_u - 1, \ldots, 0$ occurs with amplitude m_u.

spins which encode the information about the input state remain in their initial state, and we can write:

$$|\Phi_{\text{idling}}(\psi)\rangle = |\psi\rangle \otimes \frac{1}{\sqrt{L'}} \sum_{t=1}^{L'} |t\rangle$$

The rest of the history state we capture in

$$|\Phi_{\text{comp}}(\psi)\rangle \coloneqq \frac{1}{\sqrt{T-L'}} \sum_{t=L'+1}^{T} \left|\psi^{(t)}\right\rangle |t\rangle \,,$$

such that the total history state is

$$|\Phi(\psi)\rangle = \sqrt{\frac{L'}{T}} \, |\Phi_{\text{idling}}(\psi)\rangle + \sqrt{\frac{T-L'}{T}} \, |\Phi_{\text{comp}}(\psi)\rangle \,.$$

We now define the encoding $\mathcal{E}(M) = VMV^\dagger$ via the isometry

$$V = \sum_i |\Phi_{\text{idling}}(i)\rangle \langle i| \,. \tag{8}$$

where $|i\rangle$ are the computational basis states (any complete basis will suffice). $\mathcal{E}$ is a local encoding, which can be verified by a direct calculation:

$$\begin{aligned}
\mathcal{E}(A_j \otimes \mathbb{1}) &= \sum_{ik} |\Phi_{\text{idling}}(i)\rangle \langle i| \, (A_j \otimes \mathbb{1}) \, |k\rangle \langle \Phi_{\text{idling}}(k)| \\
&= \sum_{ik} |i\rangle \langle i| \, (A_j \otimes \mathbb{1}) \, |k\rangle \langle k| \otimes \frac{1}{L} \sum_{tt'=1}^{L} |t\rangle \left\langle t'\right| \\
&= (A_j \otimes \mathbb{1}) \sum_i |i\rangle \langle i| \otimes \frac{1}{L} \sum_{tt'=1}^{L} |t\rangle \left\langle t'\right| \\
&= \left(A_j^{\text{phys}} \otimes \mathbb{1}\right) \sum_i |\Phi_{\text{idling}}(i)\rangle \langle \Phi_{\text{idling}}(i)| \\
&= \left(A_j^{\text{phys}} \otimes \mathbb{1}\right) \mathcal{E}(\mathbb{1}),
\end{aligned} \tag{9}$$

where A_j^{phys} is the operator A acting on the Hilbert space corresponding to the j^{th} qudit.

We now consider the encoding $\mathcal{E}'(M) = V'MV'^\dagger$, defined via

$$V' = \sum_i |\Phi(i)\rangle \langle i| \,. \tag{10}$$

We have that

$$\begin{aligned}\|V' - V\|^2 &= \left\| \sum_i \left(|\Phi(i)\rangle \langle i| - |\Phi_{\text{idling}}(i)\rangle \langle i| \right) \right\|^2 \\ &= \left\| \sum_i \left(\sqrt{\frac{T-L'}{T}} |\Phi_{\text{comp}}(i)\rangle \langle i| + \left(\sqrt{\frac{L'}{T}} - 1 \right) |\Phi_{\text{idling}}(i)\rangle \langle i| \right) \right\|^2 \\ &\leq 2 \left(1 - \sqrt{\frac{L'}{T}} \right) \leq 2 \frac{T-L'}{T} = 2 \frac{T_{\text{PE}}}{T}.\end{aligned} \tag{11}$$

By Lemma 3.3, $S_{\mathcal{E}'}$ is the ground space of H_{SF}.

Splitting the Ground Space Degeneracy of $\mathbf{H_{SF}}$. What is left to show is that there exist one body-projectors Π_α and Π_β which add just the right amount of energy to states in the kernel $\mathcal{L}(H_{\text{SF}})$ to reproduce the target Hamiltonian's spectrum. We first choose the one body terms in H_{univ} to be projectors onto local subspaces which contain the two states which are outputs of the M_{PE} computation - $|\Omega_a\rangle$ and $|\Omega_b\rangle$:

$$\Pi_a \coloneqq \sum_{i=1}^{N} |\Omega_a\rangle\langle\Omega_a|_i \quad \text{and} \quad \Pi_b \coloneqq \sum_{i=1}^{N} |\Omega_b\rangle\langle\Omega_b|_i .$$

We have shown that if the input state is $|u\rangle$, which is an eigenstate of U with eigenphase $\phi_u = a_u\sqrt{2} - b_u$, then the history state will contain a_u terms with one spin in the state $|\Omega_a\rangle$ and b_u terms with one spin in the state $|\Omega_b\rangle$ (each term in the history state will have amplitude $\frac{1}{T}$). If the input is a general state $|\psi\rangle = \sum_u m_u |u\rangle$ then for each u the history state will contain a_u terms with one spin in the state $|\Omega_a\rangle$ and b_u terms with one spin in the state $|\Omega_b\rangle$, where now each of these terms has amplitude m_u/T.

Let $\Pi \coloneqq \sum_i |\Phi(i)\rangle \langle\Phi(i)|$ for some complete basis $|i\rangle$, and we define $H_1 \coloneqq T(\sqrt{2}\Pi_a - \Pi_b)$, where T is the total time in the computation. It thus follows that the energy of $|\Phi(u)\rangle$ with respect to the operator $\Pi H_1 \Pi$ is given by $\phi_u + \mathrm{O}(\epsilon)$.

Finally, we need the following technical lemma from [Bravyi and Hastings, 2017].

Lemma 3.6 (First-order simulation [Bravyi and Hastings, 2017])**.** *Let H_0 and H_1 be Hamiltonians acting on the same space and Π be the projector onto the ground space of H_0. Suppose that H_0 has eigenvalue 0 on Π and the next smallest eigenvalue is at least 1. Let V be an isometry such that $VV^\dagger = \Pi$ and*

$$\|V H_{\text{target}} V^\dagger - \Pi H_1 \Pi\| \leq \epsilon/2. \tag{12}$$

Let $H_{\text{sim}} = \Delta H_0 + H_1$. Then there exists an isometry $\tilde{V}$ onto the the space spanned by the eigenvectors of H_{sim} with eigenvalue less than $\Delta/2$ such that

1. $\|V - \tilde{V}\| \leq \mathrm{O}(\|H_1\|/\Delta)$

2. $\|\tilde{V} H_{\mathrm{target}} \tilde{V}^\dagger - H_{\mathrm{sim}<\Delta/2}\| \leq \epsilon/2 + \mathrm{O}(\|H_1\|^2/\Delta)$

We will apply Lemma 3.6 with $H_0 = 2T^2 H_{\mathrm{SF}}$ and $H_1 = T(\sqrt{2}\Pi_a - \Pi_b)$. We have $\lambda_{\min}(H_{\mathrm{SF}}) = 0$ and the next smallest non-zero eigenvalue of H_{SF} is $(1 - \cos(\pi/2T) \geq 1/2T^2)$ by Lemma 3.3, so $H_0 = 2T^2 H_{\mathrm{SF}}$ has next smallest non-zero eigenvalue at least 1. Moreover, $\|H_1\| = \sqrt{2}T$. Note that V', as defined in Eq. (10), is an isometry which maps onto the ground state of H_0. By construction we have that the spectrum of H_{target} is approximated to within ϵ by H_1 restricted to the ground space of H_{SF}, thus $\|\Pi H_1 \Pi - \tilde{\mathcal{E}}(H)\| \leq \epsilon$.

Lemma 3.6 therefore implies that there exists an isometry $\tilde{V}$ that maps exactly onto the low energy space of H_{univ} such that $\|\tilde{V} - V'\| \leq \mathrm{O}(\sqrt{2}T/(\Delta/2T^2)) = \mathrm{O}(T^3/\Delta)$. By the triangle inequality and Eq. (36), we have:

$$\|V - \tilde{V}\| \leq \|V - V'\| + \|V' - \tilde{V}\| \leq O\left(\frac{T^3}{\Delta} + \frac{T_{\mathrm{PE}}}{T}\right). \tag{13}$$

The second part of the lemma implies that

$$\|\tilde{V} H_{\mathrm{target}} \tilde{V}^\dagger - H_{\mathrm{univ}<\Delta'/2}\| \leq \epsilon/2 + \mathrm{O}((\sqrt{2}T)^2/(\Delta/2T^2)) = \epsilon/2 + \mathrm{O}(T^4/\Delta). \tag{14}$$

Therefore, the conditions of Definition 1.1 are satisfied for a $(\Delta', \eta, \epsilon')$-simulation of H_{target}, with $\eta = O\left(T^3/\Delta + T_{\mathrm{PE}}/T\right)$, $\epsilon' = \epsilon + \mathrm{O}(T^4/\Delta)$ and $\Delta' = \Delta/2T^2$. Therefore we must increase L so that $T \geq \mathrm{O}(T_{\mathrm{PE}}/\eta) = \mathrm{poly}(n, d^k, \|H\|, 1/\epsilon, 1/\eta)$ by Eq. (6), (thereby determining x), and increase Δ so that

$$\Delta \geq \Delta' T^2 + \frac{T^3}{\eta} + \frac{T^4}{\epsilon} \tag{15}$$

to obtain a $(\Delta', \eta, \epsilon)$-simulation of the target Hamiltonian. The claim follows. □

We can now prove our main theorem:

Theorem 3.7. *There exists a two-body interaction depending on a single parameter $h(\phi)$ such that the family of translationally-invariant Hamiltonians on a chain of length N,*

$$H_{\mathrm{univ}}(\phi, \Delta, T) = \Delta \sum_{\langle i,j \rangle} h(\phi)_{i,j} + T \sum_{i=0}^{N-1} \left(\sqrt{2}\Pi_\alpha - \Pi_\beta\right)_i, \tag{16}$$

is a universal model, where Δ, T and ϕ are parameters of the Hamiltonian, and the first sum is over adjacent site along the chain. Furthermore, the universal model is efficient in terms of the number of spins in the simulator system.

Proof. The two body interaction $h(\phi)$ makes up a standard form Hamiltonian which encodes a QTM, M_1 dovetailed with the phase-estimation computation from Lemma 3.5. The QTM M_1 carries out phase estimation on the parameter ϕ in the Hamiltonian, and writes out the binary expansion of ϕ (which contains a description of the Hamiltonian to be simulated) on its work tape. There is a standard form Hamiltonian in [Cubitt et al., 2015] which encodes this QTM, so by Lemma 3.5 we can construct a standard form Hamiltonian which simulates all quantum Hamiltonians by dovetailing M_1 with M_{PE}.

The space requirement for the computation is $\mathrm{O}(|\phi|)$, where $|\phi|$ denotes the length of the binary expansion of ϕ, and the computation requires time $T_1 = \mathrm{O}(|\phi|2^{|\phi|})$ [Cubitt et al., 2015, Theorem 10] The standard form clock construction set out in [Cubitt et al., 2015, Section 4.5] allows for computation time of $\mathrm{O}(|\phi|2^{|\phi|})$ using a Hamiltonian on $|\phi|$ spins. We therefore find that for a k-local target Hamiltonian H_{target} acting on n spins of local dimension d, the number of spins required in the simulator system for a simulation that is ϵ close to H_{target} is given by $N = \mathrm{O}(|\phi|) = \mathrm{poly}\left(n, d^k, \|H\|, 1/\eta, 1/\epsilon\right)$.

Therefore, the universal model is efficient in terms of the number of spins in the simulator system as defined in Toby's lectures. □

Note that this universal model is *not* efficient in terms of the norm $\|H_{\mathrm{univ}}\|$. This is immediately obvious, since $\|H_{\mathrm{univ}}\| = \Omega(\Delta)$, and using the relations between Δ', η, ϵ, and T and Δ from Lemma 3.5 and Eq. (15),

$$T = T_1 + L + T_{\mathrm{PE}} = O\left(2^x + \mathrm{poly}\left(n, d^k, \|H_{\mathrm{target}}\|, \frac{1}{\epsilon}, \frac{1}{\eta}\right)\right) \quad \text{and} \quad \Delta \geq \Delta' T^2 + \frac{T^3}{\eta} + \frac{T^4}{\epsilon}$$

by Eq. (6), so T, Δ are both $\mathrm{poly}\left(2^x, \|H_{\mathrm{target}}\|, \Delta', 1/\epsilon, 1/\eta\right)$. For a k-local Hamiltonian H_{target} with description x as presented in **??**, $|x| = \Omega\left(md^{2k}\log\left(\|H_{\mathrm{target}}\|md^{2k}/\delta\right)\right)$.

However if we only wish to simulate a translationally invariant k-local Hamiltonian H_{target}, this can be specified to accuracy δ with just $\log\left(\|H_{\mathrm{target}}\|md^{2k}/\delta\right)$ bits of information. In this case (for $d, k = \mathrm{O}(1)$ and taking $\delta = \epsilon$), the interaction strengths are then $\mathrm{poly}(n, \|H_{\mathrm{target}}\|, \Delta', \frac{1}{\eta}, \frac{1}{\epsilon})$, and the whole simulation is efficient.

Lemma 3.5 also allows the construction of a universal quantum simulator with two free parameters.

Theorem 3.8. *There exists a fixed two-body interaction h such that the family of translationally-invariant Hamiltonians on a chain of length N,*

$$H_{\mathrm{univ}}(\Delta, T) = \Delta \sum_{\langle i,j \rangle} h_{i,j} + T \sum_{i=0}^{N-1} \left(\sqrt{2}\Pi_\alpha - \Pi_\beta\right)_i, \tag{17}$$

is a universal model, where Δ and T are parameters of the Hamiltonian, and the first sum is over adjacent sites along the chain.

Proof. As in Theorem 3.7, the two body interaction h makes up a standard form Hamiltonian which encodes a QTM M_1 dovetailed with the phase-estimation computation from Lemma 3.5. It is based on the construction from [Gottesman and Irani, 2009b].

Take M_1 to be a binary counter Turing machine which writes out N—the length of the qudit chain—on its work tape. We will choose N to contain a description of the Hamiltonian to be simulated. There is a standard form Hamiltonian in [Gottesman and Irani, 2009b] which encodes this QTM, so by Lemma 3.5 we can construct a standard form Hamiltonian which simulates all quantum Hamiltonians by dovetailing M_1 with M_{PE}.

Since $B(N)$, as defined in Eq. (3), contains a description of the Hamiltonian to be simulated, we have that

$$N = \mathrm{poly}\left(2^{\mathrm{poly}(n,\|H_{\mathrm{target}}\|,1/\eta,1/\epsilon)}\right).$$

The standard form clock used in the construction allows for computation time polynomial in the length of the chain, so exp(poly)-time in the size of the target system. As before, by Eq. (6), we require

$$T = T_1 + L + T_{\mathrm{PE}} = O\left(N + \mathrm{poly}\left(n, d^k, \|H_{\mathrm{target}}\|, \frac{1}{\epsilon}, \frac{1}{\eta}\right)\right) \quad \text{and} \quad \Delta \geq \Delta' T^2 + \frac{T^3}{\eta} + \frac{T^4}{\epsilon}.$$

□

According to the requirements of Definition 1.1, the universal simulator of the second theorem is not efficient in either the number of spins, nor in the norm. However—as was noted in [Piddock and Bausch, 2020]—this is unavoidable if there is no free parameter in the universal Hamiltonian which encodes the description of the target Hamiltonian: a translationally invariant Hamiltonian on N spins can be described using only $\mathrm{O}(\mathrm{poly}\log(N))$ bits of information, whereas a k-local Hamiltonian which breaks translational invariance in general requires $\mathrm{poly}(N)$ bits of information. So, by a simple counting argument, we can see that it is not possible to encode all the information about a k-local Hamiltonian on n spins in a fixed translationally invariant Hamiltonian acting on $\mathrm{poly}(n)$ spins.

We observe that the parameters Δ and T are qualitatively different to ϕ, in that they do not depend on the Hamiltonian to be simulated, but only the parameters $(\Delta', \epsilon, \eta)$ determining the precision of the simulation.

4 General conditions for universality of quantum Hamiltonians

4.1 Complexity theory

Definition 4.1 ($\mathrm{QMA}(c,s)$ [Kitaev et al., 2002])**.** *A promise problem* $A = A_{\mathrm{YES}} \cup A_{\mathrm{NO}}$ *is in* $\mathrm{QMA}(c,s)$ *if and only if there exists a polynomially-bounded*

function p and a quantum polynomial time verifier V such that for all n and all $x \in \{0,1\}^n$:

- *If $x \in A_{\text{YES}}$, there exists a $p(n)$-qubit quantum state $|\psi\rangle$ such that* $\Pr[V \text{ accepts } (x, |\psi\rangle)] \geq c$
- *If $x \in A_{\text{NO}}$ for all $p(n)$-qubit quantum states $|\psi\rangle$,* $\Pr[V \text{ accepts } (x, |\psi\rangle)] \leq s$

$\text{QMA}(c, s)$ where c and s are separated by an inverse polynomial is the class QMA [Kitaev et al., 2002]. $\text{QMA}(c, s)$ where c and s are separated by an exponentially small gap is the class PreciseQMA. PreciseQMA is known to be equal to the class PSPACE [Fefferman and Lin, 2016].

Definition 4.2 ($\text{QMA}_{\text{EXP}}(c, s)$ [Gottesman and Irani, 2009a])**.** *A promise problem $A = A_{\text{YES}} \cup A_{\text{NO}}$ is in $\text{QMA}_{\text{EXP}}(c, s)$ if and only if there exists a k and a quantum exponential time verifier V such that for all n and all $x \in \{0,1\}^n$:*

- *If $x \in A_{\text{YES}}$, there exists a 2^{n^k}-qubit quantum state $|\psi\rangle$ such that* $\Pr[V \text{ accepts } (x, |\psi\rangle)] \geq c$
- *If $x \in A_{\text{NO}}$ for all 2^{n^k}-qubit quantum states $|\psi\rangle$,* $\Pr[V \text{ accepts } (x, |\psi\rangle)] \leq s$

$\text{QMA}_{\text{EXP}}(c, s)$ where c and s are separated by an inverse exponential is the class QMA_{EXP}.

The canonical problem in Hamiltonian complexity is k-LOCAL HAMILTONIAN.

k-LOCAL HAMILTONIAN (g)

Input:	Local Hamiltonian $H = \sum_{i=1}^{m} h_i$ on an N-partite Hilbert space of constant local dimension, and $m \leq \text{poly}\, N$. Each $h_i := h_{S_i} \otimes frm[o]--_{S_i^c}$ acts non-trivially on at most $\lvert S_i \rvert \leq k$ sites, and $\lVert h_i \rVert \leq 1$. Two numbers $\alpha, \beta > 0$ with $\beta - \alpha \geq g(N)$.
Promise:	The ground state energy $\lambda_{\min}(H)$ either $\geq \beta$, or $\leq \alpha$.
Question:	YES if $\lambda_{\min}(H) \leq \alpha$, else NO.

k-LOCAL HAMILTONIAN $(1/\text{poly})$ is QMA-complete for $k \geq 2$. k-LOCAL HAMILTONIAN $(1/\exp)$ is PreciseQMA-complete.

We can also consider the special case where the set of interaction terms and / or the geometry of the interaction graph is restricted (which can implicitly constrain the family's locality k).

Definition 4.3 ($\mathcal{M}$-HAMILTONIAN)**.** *The k-LOCAL HAMILTONIAN $(1/\text{poly})$ problem, where the Hamiltonian is restricted to belong to $\mathcal{M}$, some (possibly infinite) family of Hamiltonians.*

Definition 4.4 (Precise-$\mathcal{M}$-Hamiltonian)**.** *The* k-local Hamiltonian $(1/\exp)$ *problem, where the Hamiltonian is restricted to belong to* $\mathcal{M}$*, some (possibly infinite) family of Hamiltonians.*

4.2 Faithful Hamiltonian reductions

The acceptance operator, $Q(U)$, of a QMA-verification circuit, U, is defined as [Marriott and Watrous, 2005, Th. 3.6]:

$$Q(U) = \langle 0|^{\otimes m} U^\dagger \Pi_{\text{out}} U |0\rangle^{\otimes m} \tag{18}$$

where U requires m ancillas, initialised in the $|0\rangle$ state.

Definition 4.5 (Gapped acceptance operators)**.** *Consider a promise problem* A *which can be verified by a unitary circuit* U*, with completeness probability* c*. Let* $x \in A$ *with* $n = |x|$ *the size of the instance. We say the acceptance operator is gapped if it holds that* $c - \lambda_x > 1/\operatorname{poly}(n)$*, where* λ_x *is the largest eigenvalue of* $Q(U_x)$ *which is less than completeness* c*, .*

In other words, Definition 4.5 means that any state with acceptance probability below the completeness threshold already lies significantly below it, namely $1/\operatorname{poly}$ bounded away.

Note there is a subtle difference between the promise gap and the question of whether or not the acceptance operator is gapped. For any NO instance of a problem $A \in$ QMA the definition of QMA trivially implies that the acceptance operator is gapped (since the acceptance probability is below the soundness threshold $\lambda_x \leq s$, and $c - s > 1/\operatorname{poly}$ by definition). However, for YES instances it is possible to have an acceptance operator which is not gapped. We will see that for YES instances the question of whether or not the acceptance operator is gapped is related to the *spectral gap* of a Hamiltonian, rather than the promise gap.

The idea of requiring a gap in the spectrum of proof systems has arisen before in the Hamiltonian complexity literature, first in [Aharonov et al., 2008] in the definition of the class PGQMA (Polynomially Gapped QMA).[2] The notion of a gap in the spectrum of the proof system is again seen to be related to the *spectral gap* of a Hamiltonian, as k-local Hamiltonian with the added promise that the spectral gap of the Hamiltonian is inverse polynomial is complete for PGQMA [Aharonov et al., 2008].

Definition 4.6 (Faithful Hamiltonian reduction)**.** *Let* $A = A_{\text{YES}} \cup A_{\text{NO}}$ *be a promise problem which can be verified by a family of circuits,* U*, of length* T*, with completeness probability* c*. The acceptance operator* $Q(U)$ *is as defined*

[2] PGQMA is a similar class to QMA with the added condition that the acceptance operator of the verification circuit has an inverse polynomial spectral gap. Note that for PGQMA the gap is required to be between the lowest and second lowest eigenvalue, unlike in our definition.

*in Eq. (18). Consider a reduction from A to the $\mathcal{M}$-*Hamiltonian *problem. For a verification circuit with gapped $Q(U)$, we say the reduction is* faithful *with respect to U if for all instances $x \in A_{\text{YES}}$ there exists a Hamiltonian $H_x \in \mathcal{M}$ acting on $\text{poly}(n)$ qudits (where $n = |x|$) such that for the low energy subspace,*

$$\mathcal{S}_0 \coloneqq \text{span}\left\{ |\phi\rangle : H_x |\phi\rangle = \tilde{\lambda} |\phi\rangle, \tilde{\lambda} \leq \frac{\kappa(1-c)}{T+1} \right\}$$

for $\kappa = 1/\text{poly}(n)$, the following holds.

(i). $\|\Pi_{\mathcal{S}_0} - \Pi_{\mathcal{E}(\mathcal{L})}\| \leq \eta$ where

- *$\eta < 1$ can be made arbitrarily small,*
- *$\Pi_{\mathcal{S}}$ denotes the projector onto the subspace $\mathcal{S}$,*
- *$\mathcal{E}$ is some local encoding (independent of the problem being encoded),*
- *$\mathcal{L} \coloneqq \text{span}\{|\psi\rangle : Q(U)|\psi\rangle = \lambda|\psi\rangle, \lambda \geq c\}$.*

(ii). The spectral gap above the subspace $\mathcal{S}_0$ is $\Omega(1/\text{poly}(T))$.

For $x \in A_{\text{NO}}$, there are no conditions on H_x. Similarly, if $Q(U)$ is not gapped there are no conditions on the H_x.

Note that the concept of a faithful reduction is a property of a particular verification circuit,[3] not of the problem itself.

Definition 4.7. *We say that $\mathcal{M}$-*Hamiltonian *is* QMA*-complete under faithful reductions if for all $A \in$ QMA and for any polynomial time* QMA*-verification circuit U which verifies A, there exists a reduction from A to the $\mathcal{M}$-*Hamiltonian *problem which is faithful with respect to U.*

4.3 The modified Kitaev Hamiltonian

The Hamiltonian we use to prove necessity of the faithfulness condition is a modification of the 5-local Hamiltonian shown to be QMA-complete in [Kitaev et al., 2002]. Note that this choice is convenient, but the procedure we set out here to demonstrate faithfulness could be applied to any history-state Hamiltonian in the literature.

The original 5-local Hamiltonian is a "circuit-to-Hamiltonian" mapping, given by

$$H_{\text{K}} = H_{\text{in}} + H_{\text{prop}} + H_{\text{out}} + H_{\text{clock}} \tag{19}$$

where the Hamiltonian is acting on the Hilbert space

$$\mathcal{H} \coloneqq \mathcal{H}_Q \otimes \mathcal{H}_C = (\mathbb{C}^2)^{\otimes n} \otimes (\mathbb{C}^2)^{\otimes T+1} = \mathbb{C}^2 \otimes (\mathbb{C}^2)^{\otimes |W|} \otimes (\mathbb{C}^2)^{\otimes |A|} \otimes (\mathbb{C}^2)^{\otimes T} \tag{20}$$

[3] Although any equivalent model of computation could be substituted into the definition.

and

$$H_{\text{in}} = \Pi_1^{(1)} \otimes |0\rangle\langle 0|_1^c + \sum_{j \in A} \Pi_j^{(1)} \otimes |0\rangle\langle 0|_1^c \tag{21}$$

$$H_{\text{out}} = \Pi_1^{(0)} \otimes |1\rangle\langle 1|_T^c \tag{22}$$

$$H_{\text{clock}} = \mathbb{1} \otimes \sum_{t+1}^{T-1} |01\rangle\langle 01|_{t,t+1}^c \tag{23}$$

$$H_{\text{prop}} = \frac{1}{2} \sum_{t=1}^{T-1} H_t \tag{24}$$

with

$$\begin{aligned} H_t = \mathbb{1} \otimes (&|10\rangle\langle 10|_{t,t+1}^c + |10\rangle\langle 10|_{t+1,t+2}^c) \\ &- U_{t+1} \, |110\rangle\langle 110|_{t-1,t,t+1}^c - U_{t+1}^\dagger \, |100\rangle\langle 100|_{t-1,t,t+1}^c \end{aligned} \tag{25}$$

where the U_t correspond to the gates applied at time t in the circuit being encoded.

The Hamiltonian without the output penalty,

$$H_0 = H_{\text{in}} + H_{\text{prop}} + H_{\text{clock}}, \tag{26}$$

has a degenerate ground space spanned by states of the form

$$\left|\eta^{(0,\alpha)}\right\rangle = \frac{1}{\sqrt{T+1}} \sum_{t=0}^{T} \left|\gamma_t^{(0,\alpha)}\right\rangle \tag{27}$$

for arbitrary α where

$$\left|\gamma_t^{(0,\alpha)}\right\rangle = |\alpha_0(t)\rangle \otimes \left|1^t 0^{T-t}\right\rangle^c \tag{28}$$

where $|\alpha_0(t)\rangle$ is the state of the quantum circuit at time t if the input state of the ancillas and flag qubit correspond to the binary string $\mathbf{0} = 0^{1+|A|}$, and the input state of the witness is given by $|\alpha\rangle$.

The modified Kitaev Hamiltonian we use is given by

$$H_{\text{MK}} = H_{\text{in}} + H_{\text{prop}} + \kappa H_{\text{out}} + H_{\text{clock}} \tag{29}$$

where $\kappa = 1/\operatorname{poly}(T) = \mathrm{o}(1/T^3)$.

4.4 The $\mathcal{K}$-Hamiltonian problem is QMA-complete under faithful reductions

Lemma 4.8. $\mathcal{K}$-Hamiltonian *is* QMA*-complete.*

Proof. Essentially unchanged from the original proof – see [Kohler et al., 2022] for the details. □

To show that we can always choose the reduction to be faithful, we need a lemma about the spectrum and low energy subspace of H_{MK} (for a proof see [Kohler et al., 2022]).

Lemma 4.9. *Consider a modified Kitaev-Hamiltonian, H_{MK}, encoding the verification circuit of some* QMA *problem. Let $Q(U)$ be the acceptance operator for a verifier circuit U for some $A \in$ QMA. Set*

$$\mathcal{C}_0 \coloneqq \mathrm{span}\left\{\left|\eta^{(0,\phi)}\right\rangle : Q(U)\left|\phi\right\rangle = \lambda\left|\phi\right\rangle, \lambda > c\right\}$$

where c is the completeness parameter of the problem, and let $g \coloneqq c - \lambda_x$ where λ_x is the largest eigenvalue of $Q(U)$ which is less than c, as in Definition 4.5.

If $g > 2T^3(T+1)\kappa$, then there exists a unitary transformation V such that the subspace $\mathcal{S}_0$ defined by $\Pi_{\mathcal{S}_0} \coloneqq V^\dagger \Pi_{\mathcal{C}_0} V$ is the low energy subspace of H_{MK}:

$$\mathcal{S}_0 = \mathrm{span}\left\{\left|\psi\right\rangle : H_{\mathrm{MK}}\left|\psi\right\rangle = \lambda\left|\psi\right\rangle, \lambda \leq \frac{\kappa(1-c)}{T+1} + T^3\kappa^2\right\}. \tag{30}$$

$$\left\|\Pi_{\mathcal{S}_0} - \Pi_{\mathcal{C}_0}\right\| = \mathrm{O}(T^3\kappa) \tag{31}$$

and the spectral gap above $\mathcal{S}_0$ is given by $\Omega(\frac{g\kappa}{T+1} - T^3\kappa^2)$.

Lemma 4.10. *The $\mathcal{K}$-*Hamiltonian *problem is* QMA*-complete under faithful reductions.*

Proof. For any verification circuit, U, of any problem in QMA, we can require that the computation 'idles' in its initial state for L time steps before carrying out its verification computation ("idling to enhance coherence" [Aharonov et al., 2007]).

The history state of the computation for the first L time steps will be given by

$$\left|\eta^{(0,\alpha)}_{\mathrm{idling}}\right\rangle = \left|\alpha\right\rangle \otimes \left|0\right\rangle^{|A|+1} \otimes \frac{1}{\sqrt{L}} \sum_{t=1}^{L} \left|1^t 0^{T-t}\right\rangle^c. \tag{32}$$

The rest of the history state is captured in

$$\left|\eta^{(0,\alpha)}_{\mathrm{comp}}\right\rangle = \frac{1}{\sqrt{T-L}} \sum_{t=L+1}^{T} \left|\gamma^{(0,\alpha)}_t\right\rangle \tag{33}$$

So the total history state is given by

$$\left|\eta^{(0,\alpha)}\right\rangle = \sqrt{\frac{L}{T}}\left|\eta^{(0,\alpha)}_{\mathrm{idling}}\right\rangle + \sqrt{\frac{L-T}{T}}\left|\eta^{(0,\alpha)}_{\mathrm{comp}}\right\rangle \tag{34}$$

The encoding $\mathcal{E}(M) = VMV^\dagger$ defined via the isometry

$$V = \sum_i \left|\eta_{\text{idling}}^{(0,i)}\right\rangle \langle i| , \tag{35}$$

where the $|i\rangle$ are computational basis states, is local. (This can be verified by direct calculation, see [Kohler et al., 2020].)

Moreover, we have that

$$\begin{aligned}
\|\Pi_{\mathcal{C}_0} - \Pi_{\mathcal{E}(\mathcal{L})}\|^2 &= \left\| \sum_{|\phi\rangle\in\mathcal{L}} \left(\left|\eta^{(0,\phi)}\right\rangle \left\langle \eta^{0,\phi}\right| - \left|\eta_{\text{idling}}^{(0,\phi)}\right\rangle \left\langle \eta_{\text{idling}}^{(0,\phi)}\right| \right) \right\|^2 \\
&= \left\| \sum_{|\phi\rangle\in\mathcal{L}} \left(\sqrt{\frac{T-L}{T}} \left|\eta_{\text{comp}}^{(0,\phi)}\right\rangle \left\langle \eta_{\text{comp}}^{(0,\phi)}\right| + \left(\sqrt{\frac{L}{T}} - 1\right) \left|\eta_{\text{idling}}^{(0,\phi)}\right\rangle \left\langle \eta_{\text{idling}}^{(0,\phi)}\right| \right) \right\|^2 \\
&\leq 2\left(1 - \sqrt{\frac{L}{T}}\right)
\end{aligned} \tag{36}$$

where $\mathcal{L} \coloneqq \{|\phi\rangle : Q|\phi\rangle = \lambda|\phi\rangle, \lambda > c\}$ Therefore $\|\Pi_{\mathcal{C}_0} - \Pi_{\mathcal{E}(\mathcal{L})}\|$ can be made arbitrarily small by increasing L.

The result follows immediately from Lemma 4.9 and the triangle inequality. □

4.5 General conditions for universality of quantum Hamiltonians

In order to state our main theorem we require one more definition.

Definition 4.11 (Closed Hamiltonian model)**.** *We say that a model $\mathcal{M}$, is closed if for any pair of Hamiltonians $H_A^{(1)}, H_B^{(2)} \in \mathcal{M}$ acting on sets of qudits A, B respectively where in general $A \cap B \neq \{\}$, there exists a Hamiltonian $H^{(3)} \in \mathcal{M}$ which can efficiently simulate $H_A^{(1)} + H_B^{(2)}$.*

We can now prove our main result:

Theorem 4.12 (Universality Classification)**.** *A family of Hamiltonians, $\mathcal{M}$, is an efficient universal model iff $\mathcal{M}$-*Hamiltonian *is* QMA*-complete under faithful reductions, and $\mathcal{M}$ is closed.*

Proof. First consider the only if direction. Closure is clearly necessary: if a model $\mathcal{M}$ is universal, all Hamiltonians (including those of the form $H_A^{(1)} + H_B^{(2)}$ for $H_A^{(1)}, H_B^{(2)} \in \mathcal{M}$) can be simulated by a Hamiltonian in the model. In Lemma 4.10 we proved the $\mathcal{K}$-Hamiltonian problem is QMA-complete under faithful reductions. Any efficient universal model must be

able to simulate Hamiltonians in $\mathcal{K}$ with only polynomial overhead, hence $\mathcal{M}$-Hamiltonian must itself be QMA-complete under faithful reductions.

Now consider the if direction. Let $\mathcal{M}$ be a family of Hamiltonians meeting the conditions of the theorem, i.e. such that $\mathcal{M}$-Hamiltonian is QMA-complete under faithful reductions, and $\mathcal{M}$ is closed. We will explicitly construct a universal model, solely based on these conditions.

Consider the following problem:

Yes-Hamiltonian

Input: A k-local Hamiltonian H_{target} acting on n spins with local dimension d.

Question: Output YES

This problem is (clearly) trivial. But we can choose to construct a non-trivial QMA verification circuit for it. We will choose a verification circuit which picks out a particular subspace that allows us to prove universality. By Definition 4.7 there must be a faithful reduction with respect to *this* verification circuit from Yes-Hamiltonian to $\mathcal{M}$-Hamiltonian.

The verification circuit we choose, and the subspace it picks out, are captured in the following.

Lemma 4.13. Yes-Hamiltonian *can be verified by a circuit* U_a *with gapped acceptance operator* $Q(U_a)$ *with ground space*

$$\mathcal{L}_0 \coloneqq \text{span}\{|\phi\rangle : Q(U_a)\,|\phi\rangle = |\phi\rangle\} \tag{37}$$

satisfying

$$\|\Pi_{\mathcal{L}_0} - \Pi_{\mathcal{W}}\| \leq \mathrm{O}\left(a^{-1}\right) \tag{38}$$

where

$$\mathcal{W} \coloneqq \text{span}\left\{|w_\mu\rangle = \frac{1}{\sqrt{a^2+1}}\,|\psi_\mu\rangle\,(a\,|\#\rangle + |E_\mu\rangle) : H_{\text{target}}\,|\psi_\mu\rangle = E_\mu\,|\psi_\mu\rangle\right\}. \tag{39}$$

Proof. The verifier circuit, C_V, acts on the witness and two ancilla registers, B, B'. It will be helpful to divide the witness into two separate registers: An A register, which is n d-dimensional qudits. And an A' register, which consists of m qutrits with orthonormal basis states $|\#\rangle$, $|0\rangle$ and $|1\rangle$, where $m = \log_2(\epsilon)$. The B register is the same size as the A' register. The B' register consists of a single qubit.

The verifier C_V operates as follows:

1. Apply a unitary rotation $P_a : |0\rangle \to \frac{1}{\sqrt{a^2+1}}(a\,|\#\rangle + |1\rangle)$ to the B' register.f

2. Carry out controlled-phase-estimation on the A register with respect to the unitary generated by H_{target}, $U = e^{iH_{\text{target}}\tau}$, for some τ such that

$\|H_{\text{target}}\tau\| < 2\pi$. The B' register serves as the control qubit. Calculate (an approximation to) the energy E_μ from the eigenphase θ_μ and store the result in the B register (in binary).

Calculating E_μ to accuracy ϵ requires calculating the eigenphase θ_μ to accuracy $\mathrm{O}(\epsilon/\|H_{\text{target}}\|)$ which takes $\mathrm{O}(\|H_{\text{target}}\|/\epsilon)$ uses of $U = e^{iH_{\text{target}}\tau}$. The unitary U must therefore be implemented to accuracy $O(\epsilon/\|H_{\text{target}}\|)$, which can be done with overhead $\mathrm{poly}(n, d^k, \|H_{\text{target}}\|, \tau, 1/\epsilon)$ where n is system size, d is local dimension and k is locality via [Kohler et al., 2020, Lemma 3.3]. The whole procedure takes time $T_{\text{PE}} = \mathrm{poly}(n, d^k, \|H_{\text{target}}\|/\epsilon)$

3. Carry out a SWAP test [Buhrman et al., 2001] between registers A' and B. Accept if outcome 0 is measured, reject otherwise.

The entire procedure takes time $T = \mathrm{O}(\mathrm{poly}(n, d^k, \|H_{\text{target}}\|)/\epsilon)$.

Let

$$|\alpha_\mu\rangle = \sum_j \left(\frac{1}{2^m} \sum_{k=0}^{2^m-1} e^{2\pi i k(E_\mu - j/2^m)} \right) |j\rangle \tag{40}$$

be the result of applying the phase estimation algorithm on $|\psi_\mu\rangle$ with respect to $U = e^{iH_{\text{target}}\tau}$. Then evidently

$$|\phi_\mu\rangle = \frac{1}{\sqrt{a^2+1}} |\psi_\mu\rangle_A \left(a\,|\#\rangle_{A'} + |\alpha_\mu\rangle_{A'}\right) \tag{41}$$

is an eigenvector of $Q(U_a)$ with eigenvalue 1, and all eigenvectors of $Q(U_a)$ with eigenvalue 1 are in $\mathrm{span}\{|\phi_\mu\rangle\}$.

Moreover,

$$\langle w_\mu|\phi_\mu\rangle \leq \frac{a^2 + \frac{4}{\pi^2}}{a^2+1} \tag{42}$$

Therefore

$$\| \, |w_\mu\rangle\langle w_\mu| - |\phi_\mu\rangle\langle\phi_\mu| \, \| \leq 2\sqrt{1 - \left(\frac{a^2 + \frac{4}{\pi^2}}{a^2+1}\right)^2} \tag{43}$$

and

$$\|\Pi_{\mathcal{L}_0} - \Pi_{\mathcal{W}}\| \leq \mathrm{O}\left(a^{-1}\right). \tag{44}$$

The next largest eigenvalue of $Q(U_a)$ is $\frac{1}{2}$. □

It follows immediately from Eq. (44) that

$$\|\mathcal{E}(\Pi_{\mathcal{L}_0}) - \mathcal{E}(\Pi_{\mathcal{W}})\| \leq \mathrm{O}\left(a^{-1}\right) \tag{45}$$

for any encoding $\mathcal{E}$.

It follows from the triangle inequality and Definition 4.6(i) that for any instance of YES-HAMILTONIAN there exists $H_{\mathrm{LS}} \in \mathcal{M}$ with low energy subspace $\mathcal{S}_0 \coloneqq \mathrm{span}\{|\phi\rangle : H_x |\phi\rangle = \lambda_{\min} |\phi\rangle\}$ such that

$$\|\Pi_{\mathcal{S}_0} - \Pi_{\mathcal{E}(\mathcal{W})}\| \leq \eta + \mathrm{O}\left(a^{-1}\right) \tag{46}$$

where $\mathcal{E} = V\left(M \otimes P + \overline{M} \otimes Q\right) V^\dagger$ is some local encoding and η can be chosen to be arbitrarily small. The spectral gap above $\mathcal{S}_0$ is $\Omega(1/\operatorname{poly}(T))$.

Another trivial problem (that is therefore also evidently in QMA) is:

FLAG IDENTIFICATION
Input: Classical description of a one-qudit state $|f\rangle$
Question: Output YES.

For this problem we will use a faithful reduction with respect to the non-trivial verification circuit which simply measures a single qudit in the f basis. So, for any single qudit state $|f\rangle$, there exists $H_f \in \mathcal{M}$ such that:

- $H_f |\mathcal{E}_{\mathrm{state}}(\phi)\rangle = \lambda_0^{(f)} |\mathcal{E}_{\mathrm{state}}(\phi)\rangle$ for all $|\phi\rangle$ such that $\langle\phi|f\rangle = 0$
- $H_f |\mathcal{E}_{\mathrm{state}}(f)\rangle = \lambda_1^{(f)} |\mathcal{E}_{\mathrm{state}}(f)\rangle$

for some local encoding $\mathcal{E}_{\mathrm{state}}$, where $\lambda_k^{(f)}$ can be efficiently computed. (Since the problem size is $O(1)$ for a state $|f\rangle$ that can be described in $O(1)$ bits.) Wlog we will take $\lambda_k^{(f)} = k$.

Consider a Hamiltonian acting on N spins:

$$H_{\mathrm{sim}} = \Delta\left(H_{\mathrm{LS}} - \lambda_{\min} \mathbb{1}\right) + a \sum_{i=n'+1}^{N} 2^{i-(n'+1)} H_i^{(1)} \tag{47}$$

Where $H_{\mathrm{LS}} \in \mathcal{M}$ is a faithful reduction (with respect to the verifier defined in Lemma 4.13) from YES-HAMILTONIAN for the Hamiltonian $H_{\mathrm{target}} = \sum_\mu E_\mu |\psi_\mu\rangle\langle\psi_\mu|$. E_μ in $|w_\mu\rangle$ is expressed in binary to precision ϵ in qudits $[n'+1, N]$.[4] The $H_i^{(1)} \in \mathcal{M}$ are faithful reductions (with respect to the obvious verification circuit) from FLAG IDENTIFICATION for the flag states $|1\rangle$ acting on the i^{th} qudit. We will require $\Delta > \|H_{\mathrm{target}}\|$.

First we show that H_{sim} can simulate $H' = \sum_\mu E_\mu |w_\mu\rangle\langle w_\mu|$. The low energy subspace of H_{LS} consists of states in the subspace $\mathcal{S}_0$. On states in $\mathcal{S}_0$, $H_{\mathrm{LS}} - \lambda_{\min}$ has energy zero. While on states in $\mathcal{E}(\mathcal{W})$, $a \sum_{i=n'+1}^{N} 2^{i-(n'+1)} H_i^{(1)}$ has energy in the range $[E_\mu - \epsilon, E_\mu + \epsilon]$.

It follows from [Cubitt et al., 2018, Lemma 24] and Eq. (46) that there exists an encoding $\mathcal{E}'(M) = V'\left(M \otimes P + \overline{M} \otimes Q\right) V'^\dagger$ such that

$$\|V' - V\| \leq \sqrt{2}(\eta + O(a^{-1})) \tag{48}$$

[4] Here n' is the number of spins in the encoded witness state $\mathcal{E}(|w_\mu\rangle)$. Since this is a reduction to QMA we have $n' = \mathrm{O}(\operatorname{poly}(n))$.

and $\mathcal{E}'(\mathbb{1}) = \Pi_{\mathcal{S}_0}$. Moreover,

$$V' = WV \tag{49}$$

where W is a unitary satisfying

$$\Pi_{\mathcal{S}_0} = W\Pi_{\mathcal{E}(\mathcal{W})}W^\dagger \tag{50}$$

and

$$\|W - \mathbb{1}\| \leq \sqrt{2}\|\Pi_{\mathcal{S}_0} - \Pi_{\mathcal{E}(\mathcal{W})}\| \leq \mathrm{O}(\eta + a^{-1}) \tag{51}$$

We need the following technical lemma.

Lemma 4.14 (First-order simulation [Bravyi and Hastings, 2017][Lemma 14)**.** *] Let H_0 and H_1 be Hamiltonians acting on the same space and Π be the projector onto the ground space of H_0. Suppose that H_0 is zero on Π and the next smallest eigenvalue is at least 1. Let U be an isometry such that $UU^\dagger = \Pi$ and*

$$\|UH_{\mathrm{target}}U^\dagger - \Pi H_1 \Pi\| \leq \epsilon/2. \tag{52}$$

Let $H_{\mathrm{sim}} = \Delta H_0 + H_1$. Then there exists an isometry $\tilde{V}$ onto the the space spanned by the eigenvectors of H_{sim} with eigenvalue less than $\Delta/2$ such that

1. $\|U - \tilde{V}\| \leq \mathrm{O}\left(\Delta^{-1}\|H_1\|\right)$
2. $\|\tilde{V}H_{\mathrm{target}}\tilde{V}^\dagger - H_{\mathrm{sim}<\Delta/2}\| \leq \mathrm{O}\left(\Delta^{-1}\|H_1\|^2\right) + \epsilon/2$

We will apply Lemma 4.14 with $H_1 = a\sum_{i=n'+1}^{N} 2^{i-(n'+1)}H_i^{(1)}$ and $H_0 = \delta H_{\mathrm{LS}}$ where $\delta = \mathrm{O}(\mathrm{poly}(T))$. We have that in $\Pi_{\mathcal{S}_0}$, H_{LS} has energy zero and by Definition 4.6(ii) the spectral gap above $\mathcal{S}_0$ scales as $\Omega(1/\mathrm{poly}(T))$ so $H_0 = \delta H_{\mathrm{LS}}$ has next smallest eigenvalue at least 1.

Moreover, $\|H_1\| = a\|H_{\mathrm{target}}\|$. Note that V' is an isometry which maps onto the ground state of H_0, $\mathcal{S}_0$. By construction we have that the spectrum of H_{target} is approximated to within ϵ by H_1 restricted to $\mathcal{E}(\mathcal{W})$, so $\|\Pi_{\mathcal{E}(\mathcal{W})}H_1\Pi_{\mathcal{E}(\mathcal{W})} - \mathcal{E}(H_{\mathrm{target}})\| \leq \epsilon$.

Using that the operator norm is unitarily invariant, and that $V' = WV$ gives

$$\|W\Pi_{\mathcal{E}(\mathcal{W})}H_1\Pi_{\mathcal{E}(\mathcal{W})}W^\dagger - \mathcal{E}'(H_{\mathrm{target}})\| \leq \epsilon. \tag{53}$$

We also have

$$\begin{aligned}\|\Pi_{\mathcal{S}_0}H_1\Pi_{\mathcal{S}_0} - W\Pi_{\mathcal{E}(\mathcal{W})}H_1\Pi_{\mathcal{E}(\mathcal{W})}W^\dagger\| &= \|\Pi_{\mathcal{S}_0}H_1\Pi_{\mathcal{S}_0} - \Pi_{\mathcal{S}_0}WH_1W^\dagger\Pi_{\mathcal{S}_0}\| \\ &\leq \|H_1 - WH_1W^\dagger\| \\ &\leq 2\|H_1\|\|\mathbb{1} - W\| \\ &\leq \mathrm{O}\left(a\eta\|H_{\mathrm{target}}\|\right)\end{aligned} \tag{54}$$

where we have used [Cubitt et al., 2018, Lemma 18] in the penultimate step. So

$$\|\Pi_{\mathcal{S}_0}H_1\Pi_{\mathcal{S}_0} - \mathcal{E}'(H_{\mathrm{target}})\| \leq \epsilon + \mathrm{O}\left(a\eta\|H_{\mathrm{target}}\|\right). \tag{55}$$

Lemma 4.14 therefore implies that there exists an isometry $\tilde{V}$ that maps exactly onto the low energy space of H_{sim} such that $\|\tilde{V} - V'\| \leq \mathrm{O}\left(\|H_{\text{target}}\| a/(\Delta/\delta)\right) = \mathrm{O}\left(a\delta\|H_{\text{target}}\|/\Delta\right)$. By the triangle inequality and Eq. (48), we have

$$\|V - \tilde{V}\| \leq \|V - V'\| + \|V' - \tilde{V}\| \leq \mathrm{O}\left(\frac{a\,\mathrm{poly}(T')\|H_{\text{target}}\|}{\Delta} + \eta + a^{-1}\right). \quad (56)$$

The second part of the lemma implies that

$$\|\tilde{V}H'\tilde{V}^\dagger - (H_{\text{sim}})_{<\Delta'/2}\| \leq \epsilon + \mathrm{O}\left(a\eta\|H_{\text{target}}\| + (a|H_{\text{target}}\|)^2/(\Delta/\delta)\right) \quad (57)$$

$$= \epsilon + \mathrm{O}\left(a\eta\|H_{\text{target}}\| + \frac{a^2|H_{\text{target}}\|^2\delta}{\Delta}\right). \quad (58)$$

Therefore, the conditions of Definition 1.1 are satisfied for a $(\Delta', \eta', \epsilon')$-simulation of H', with $\eta' = \mathrm{O}\left(\frac{a\,\mathrm{poly}(T')\|H_{\text{target}}\|}{\Delta} + \eta + a^{-1}\right)$,
$\epsilon' = \epsilon + \mathrm{O}\left(a\eta\|H_{\text{target}}\| + \frac{a^2\,\mathrm{poly}(T')\|H_{\text{target}}\|^2}{\Delta}\right)$ and $\Delta' = \Delta/\delta = \Delta/\,\mathrm{poly}(T)$.

By definition we can choose η to be arbitrarily small. We can also make $\mathrm{O}\left(a^{-1}\right)$ arbitrarily small. By increasing T, we can also make ϵ arbitrarily small. Therefore, by choosing Δ such that

$$\Delta \geq \Delta'\,\mathrm{poly}(T') + \frac{a\,\mathrm{poly}(T')\|H_{\text{target}}\|}{\eta'} + \frac{a^2\,\mathrm{poly}(T')\|H_{\text{target}}\|^2}{\epsilon'} \quad (59)$$

we can construct H_{sim} which is a $(\Delta', \eta', \epsilon')$-simulation of H' with arbitrarily small ϵ', η'. Since H_{sim} is a sum of Hamiltonians which are all in $\mathcal{M}$, by the closure property there exists $H_{\text{univ}} \in \mathcal{M}$ which can efficiently simulate H_{sim}. Therefore, since simulations compose [Cubitt et al., 2018, Lemma 17] H_{univ} can simulate H'.

Finally, we show that $H' = \sum_\mu E_\mu |w_\mu\rangle\langle w_\mu|$ is itself a simulation of H_{target}. Consider the local encoding

$$\mathcal{E}'(M) = WMW^\dagger, \quad (60)$$

where $W = \sum_\mu |\psi_\mu\rangle |0\rangle \langle\psi_\mu|$, and the non local encoding

$$\tilde{\mathcal{E}}'(M) = \tilde{W}M\tilde{W}^\dagger \quad (61)$$

withk

$$\tilde{W} = \frac{1}{\sqrt{a^2+1}} \sum_\mu |\psi_\mu\rangle \left(a|\#\rangle + |E_\mu\rangle\right) \langle\psi_\mu|. \quad (62)$$

We have that

$$\|W - \tilde{W}\| = 2\left(1 - \frac{a}{\sqrt{a^2+1}}\right) \quad (63)$$

So by increasing a we can make the norm arbitrarily small. We also have that $S_{\tilde{\mathcal{E}}'} = S_{H'}$, so condition i from Definition 1.1 is met. The spectrum of H' is exactly the spectrum of $\tilde{\mathcal{E}}'(H_{\text{target}})$, so condition ii of Definition 1.1 is also met. Therefore H' is a simulation of H_{target}.

Using the composition of simulations again, we have that H_{univ} can simulate H_{target}. We have left H_{target} arbitrary, so $\mathcal{M}$ is a universal model.

Finally we consider efficiency. The simulation of H_{target} by H' is clearly efficient. To see that the simulation of H' by H_{sim} is efficient, note that the number of qudits in the simulation, N, must be polynomial in n and $\|H_{\text{target}}\|$ as H_{LS} is in QMA. Furthermore, $\|H_{\text{sim}}\| = \Omega(\Delta) = \text{poly}(T', \|H_{\text{target}}\|, 1/\epsilon', 1/\eta') = \text{poly}(n, \|H_{\text{target}}\|, 1/\epsilon', 1/\eta')$. Thus H_{sim} is an efficient simulation of H'. □

It may seem od

4.6 General conditions for universality of succinct and precise Hamiltonians

There are two corollaries about universal Hamiltonians which aren't efficient in the sense of [Cubitt et al., 2018], but which are nonetheless interesting universal models which are better suited to some applications.

Definition 4.15. *We say a Hamiltonian can be described succinctly if it can be described by* $\text{O}(\log(n))$ *bits of information when acting on* n *qudits.*

Corollary 4.16. *Let* $\mathcal{M}$ *be a family of succinct Hamiltonians. Then* $\mathcal{M}$ *is universal and can efficiently simulate any succinct Hamiltonian iff* $\mathcal{M}$-Hamiltonian *is* QMA_{EXP}*-complete under faithful reductions and* $\mathcal{M}$ *is closed.*

Proof. This proof relies on the same ideas as Theorem 4.12, so here we sketch the main ideas, highlighting where the proofs differ. First, note the proof that $\mathcal{K}$-Hamiltonian is QMA-complete under faithful reductions (Lemma 4.10) can be repurposed to prove that the $\mathcal{M}$-Hamiltonian which was shown to be QMA_{EXP}-complete in [Gottesman and Irani, 2009b] remains QMA_{EXP} complete under faithful reductions. The only if direction follows immediately as in Theorem 4.12 (where now the overhead can be exponential).

To see the if direction, we will assume there exists a family of Hamiltonians, $\mathcal{M}$, such that $\mathcal{M}$-Hamiltonian is QMA_{EXP}-complete under faithful reductions and $\mathcal{M}$ is closed. Consider the following computational problem:

Succinct-Yes-Hamiltonian

Input: A k-local Hamiltonian H_{target} acting on n spins with local dimension d, which can be described succinctly.

Question: Output YES

Succinct-Yes-Hamiltonian is clearly a trivial problem. But, as with Yes-Hamiltonian, we can construct a non-trivial verification circuit, which

picks out a particular subspace that allows us to prove universality. By Definition 4.7 there must be a faithful reduction with respect to *this* verification circuit from Succinct-Yes-Hamiltonian to $\mathcal{M}$-Hamiltonian.

The verification circuit we choose, and the subspace it picks out, are as in Lemma 4.13. Although the circuit is unchanged, this is now a $\mathrm{QMA}_{\mathrm{EXP}}$-verification circuit, since for Hamiltonians acting on n qudits the circuit length and witness size are of order $\mathrm{poly}(n) = \mathrm{O}(2^{\mathrm{poly}(x)})$, where x is the number of bits of information needed to describe the input to the problem.

It then follows, using the same argument as in Lemma 4.13, that for any Hamiltonian H_{target} acting on n spins which can be described succinctly, there exists a Hamiltonian in $\mathcal{M}$ which can simulate H_{target} efficiently (where efficiency is defined in terms of number of qudits, not bits of information).

To prove universality, note that in [Kohler et al., 2020, Theorem 3.6] a construction is given of a universal Hamiltonian, H_{succ}, (with exponential overhead in terms of number of spins and norm of simulating system) which can be described succinctly.

Since there exists a Hamiltonian in $\mathcal{M}$ which can simulate H_{succ} (for any values of the parameters in H_{succ}), and since simulations compose, it follows that $\mathcal{M}$ is a universal model. When simulating general (non-succinct) Hamiltonians, the universal model, $\mathcal{M}$, inherits an exponential overhead in terms of the numbers of qudits and the norm of the simulating system from H_{succ}. □

$\mathrm{QMA}_{\mathrm{EXP}}$ is a more powerful complexity class than QMA so it may seem odd that it appears to be less efficient as a simulator. However, there are some situations where using a family of Hamiltonians meeting the conditions of Corollary 4.16 will give a more efficient simulator than using a family of Hamiltonians meeting the conditions of Theorem 4.12. To see this, note that given a Hamiltonian which can be described succinctly, it can be simulated efficiently (in the sense of [Cubitt et al., 2018] i.e. in terms of numbers of qudits, simulating system norm, and ϵ and η parameters) by either a family of Hamiltonians with a QMA-complete $\mathcal{M}$-Hamiltonian, or a family of Hamiltonians with a $\mathrm{QMA}_{\mathrm{EXP}}$-complete $\mathcal{M}$-Hamiltonian. However, the simulation by the QMA-complete family of Hamiltonians will not be efficient in terms of the number of bits needed to describe the simulating Hamiltonian. Whereas the simulation by the $\mathrm{QMA}_{\mathrm{EXP}}$-complete family of Hamiltonians would be. So there are situations where simulation using a $\mathrm{QMA}_{\mathrm{EXP}}$-complete family of Hamiltonians is more efficient, demonstrating that the question of which family of Hamiltonians is a 'more powerful' simulator doesn't have a straightforward answer.

The obvious example of Hamiltonians that can be succinctly described are translationally invariant Hamiltonians. Examples of translationally invariant universal Hamiltonians are constructed in [Piddock and Bausch, 2020, Kohler et al., 2020], where it is noted that a translationally invariant uni-

versal model with fixed interactions must have an exponential overhead in terms of number of spins by a simple counting argument.

By considering the problem PRECISE-$\mathcal{M}$-HAMILTONIAN and introducing the idea of a *exponentially faithful*, reduction we can also derive conditions for universal models which are efficient in terms of the number of qudits in the simulator system, but not in terms of the simulating system's norm. We say a reduction is *exponentially faithful* if it meets the conditions of Definition 4.6, but where now the gap in the spectrum of the acceptance operator, g, and the corresponding gap in the spectrum of the Hamiltonian is required to satisfy $g > 1/\exp(n)$ for n the size of the input. This is a natural relaxation when considering PRECISE-$\mathcal{M}$-HAMILTONIAN, as it requires the gap in the spectrum of the acceptance operator in YES cases to scale in the same way as the promise gap for the problem.

The natural complexity class when considering PRECISE-$\mathcal{M}$-HAMILTONIAN is PreciseQMA. It is known that every problem in PreciseQMA can be solved by a quantum circuit of length $T = \mathrm{O}(\exp(n))$ acting on $\mathrm{poly}(n)$ qudits, with completeness $c = 1 - 2^{\mathrm{poly}(n)}$ and soundness $s = 2^{\mathrm{poly}(n)}$ (where n is the size of the problem input) [Fefferman and Lin, 2016, Corollary 10]. Therefore, when defining what it means for PRECISE-$\mathcal{M}$-HAMILTONIAN to be PreciseQMA-complete under exponentially faithful reductions there are two classes of circuits we could require faithfulness with respect to - the polynomial sized circuits that give exponentially small completeness-soundness gap, or the exponential sized circuits that give completeness-soundness gap exponentially close to 1. Here we choose the latter, and define:

Definition 4.17. *We say that* PRECISE-$\mathcal{M}$-HAMILTONIAN *is* PreciseQMA*-complete under exponentially faithful reductions if for all* $A \in$ PreciseQMA *and for any exponential time verification circuit* U *which verifies* A*, there exists a reduction from* A *to the* PRECISE-$\mathcal{M}$-HAMILTONIAN *problem which is exponentially faithful with respect to* U*.*

Corollary 4.18. *A family of Hamiltonians,* $\mathcal{M}$*, is a universal model which is*

1. *efficient in terms of the numbers of qudits,*
2. *not efficient in terms of the simulating system's norm and*
3. *achieves exponential accuracy in the* ϵ *parameter with polynomial overhead in number of qudits and exponential overhead in simulating system norm*

iff

i. PRECISE-$\mathcal{M}$-HAMILTONIAN *is* PSPACE*-complete under exponentially faithful reductions,*

ii. $\mathcal{M}$ *is closed and*

iii. $\mathcal{M}$-HAMILTONIAN *is not* QMA*-complete under faithful reductions.*

Proof. Recall that PSPACE = PreciseQMA [Fefferman and Lin, 2016].

First consider the if direction. Assume we have a family of Hamiltonians $\mathcal{M}$ meeting the conditions $i - iii$ of the theorem. Since $\mathcal{M}$-HAMILTONIAN is not QMA-complete under faithful reductions we have to use exponentially faithful reductions to PRECISE-$\mathcal{M}$-HAMILTONIAN, where by Definition 4.17 we are considering faithfulness with respect to exponential time circuits. The if direction follows immediately since going through the proof of Theorem 4.12 with an exponentially small gap in the acceptance operator and exponentially long computation time requires an exponentially large energy penalty Δ, and gives an exponentially small accuracy parameter ϵ.

To see the only if direction, consider a universal model, $\mathcal{M}$ meeting conditions $1 - 3$ of the theorem. Necessity of closure is trivial. Note, $\mathcal{M}$-HAMILTONIAN cannot be QMA-complete under faithful reductions, because by Theorem 4.12 if it was the universal model would be efficient in terms of norm.

Consider simulating the family of Hamiltonians $\mathcal{K}$ using the model $\mathcal{M}$ to exponential accuracy in ϵ. This demonstrates that PRECISE-$\mathcal{M}$-HAMILTONIAN is PSPACE-complete (including under exponentially faithful reductions), but it does not contradict the statement that $\mathcal{M}$-HAMILTONIAN is not QMA-complete. This is because the k-LOCAL HAMILTONIAN problem requires that the terms in the Hamiltonian are of order 1, which requires dividing each term in the simulator system by the simulating system norm, which by assumption is exponential in the size of the system. This leads to a Hamiltonian with an exponentially small spectral gap, which attenuates the promise gap too fast to maintain QMA-completeness, but gives PreciseQMA-completeness (and therefore PSPACE-completeness).

□

An example of a universal Hamiltonian meeting the conditions of Corollary 4.18 is given in [Kohler et al., 2020]. It is a translationally invariant universal model, but includes a phase parameter which encodes information about the target system, so the interactions are not fixed.

References

[Aharonov et al., 2008] Aharonov, D., Ben-Or, M., Brandao, F. G. S. L., and Sattath, O. (2008). The pursuit for uniqueness: Extending valiant-vazirani theorem to the probabilistic and quantum settings.

[Aharonov et al., 2007] Aharonov, D., van Dam, W., Kempe, J., Landau, Z., Lloyd, S., and Regev, O. (2007). Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM Journal of Computing*, 37(1):166–194.

[Bausch et al., 2018] Bausch, J., Cubitt, T., Lucia, A., and Perez-Garcia, D. (2018). Undecidability of the Spectral Gap in One Dimension.

[Bausch et al., 2017] Bausch, J., Cubitt, T., and Ozols, M. (2017). The Complexity of Translationally Invariant Spin Chains with Low Local Dimension. *Annales Henri Poincaré*, 18(11):3449–3513.

[Bernstein and Vazirani, 1997] Bernstein, E. and Vazirani, U. (1997). Quantum complexity theory. *SIAM Journal on Computing*, 26(5):1411–1473.

[Berry et al., 2005] Berry, D. W., Ahokas, G., Cleve, R., and Sanders, B. C. (2005). Efficient quantum algorithms for simulating sparse Hamiltonians.

[Berry et al., 2015] Berry, D. W., Childs, A. M., and Kothari, R. (2015). Hamiltonian simulation with nearly optimal dependence on all parameters.

[Bravyi and Hastings, 2017] Bravyi, S. and Hastings, M. (2017). On complexity of the quantum Ising model. *Communications in Mathematical Physics*, 349(1):1–45.

[Buhrman et al., 2001] Buhrman, H., Cleve, R., Watrous, J., and de Wolf, R. (2001). Quantum fingerprinting. *Physical Review Letters*, 87(16).

[Cubitt et al., 2018] Cubitt, T., Montanaro, A., and Piddock, S. (2018). Universal Quantum Hamiltonians. *Proceedings of the National Academy of Sciences*, 115(38):9497–9502.

[Cubitt et al., 2015] Cubitt, T. S., Perez-Garcia, D., and Wolf, M. M. (2015). Undecidability of the spectral gap. *Nature*, 528(7581):207–211.

[De las Cuevas and Cubitt, 2016] De las Cuevas, G. and Cubitt, T. S. (2016). Simple universal models capture all classical spin physics. *Science*, 351(6278):1180–1183.

[Fefferman and Lin, 2016] Fefferman, B. and Lin, C. Y.-Y. (2016). A Complete Characterization of Unitary Quantum Space. In *Leibniz International Proceedings in Informatics (LIPIcs)*. 9th Innovations in Theoretical Computer Science Conference.

[Feynman, 1985] Feynman, R. P. (1985). Quantum Mechanical Computers. *Optics News*, 11(2):11.

[González-Guillén and Cubitt, 2018] González-Guillén, C. E. and Cubitt, T. S. (2018). History-state hamiltonians are critical.

[Gottesman and Irani, 2009a] Gottesman, D. and Irani, S. (2009a). The quantum and classical complexity of translationally invariant tiling and hamiltonian problems.

[Gottesman and Irani, 2009b] Gottesman, D. and Irani, S. (2009b). The Quantum and Classical Complexity of Translationally Invariant Tiling and Hamiltonian Problems. *Theory of Computing*, 9(1):31–116.

[Kitaev et al., 2002] Kitaev, A. Y., Shen, A., and Vyalyi, M. N. (2002). Classical and quantum computing. In *Quantum Information*, pages 203–217. Springer New York, New York, NY.

[Kohler et al., 2020] Kohler, T., Piddock, S., Bausch, J., and Cubitt, T. (2020). Translationally-invariant universal quantum hamiltonians in 1d.

[Kohler et al., 2022] Kohler, T., Piddock, S., Bausch, J., and Cubitt, T. (2022). General conditions for universality of quantum hamiltonians. *PRX Quantum*, 3(1).

[Lloyd, 1996] Lloyd, S. (1996). Universal Quantum Simulators. *Science*, 273(5278):1073–1078.

[Marriott and Watrous, 2005] Marriott, C. and Watrous, J. (2005). Quantum arthur-merlin games.

[Piddock and Bausch, 2020] Piddock, S. and Bausch, J. (2020). Universal translationally-invariant hamiltonians.

[Usher et al., 2017] Usher, N., Hoban, M. J., and Browne, D. E. (2017). Nonunitary quantum computation in the ground space of local Hamiltonians. *Physical Review A*, 96(3):032321.

[Watson, 2019] Watson, J. D. (2019). Detailed analysis of circuit-to-hamiltonian mappings.