Universal quantum Hamiltonians

Supporting Information

Toby Cubitt[1], Ashley Montanaro[2] and Stephen Piddock[2]

[1] *Department of Computer Science, University College London, UK*
[2] *School of Mathematics, University of Bristol, UK.*

June 22, 2018

Contents

www.pnas.org/cgi/doi/10.1073/pnas.1804949115

1 Introduction

In this supplementary material we start by introducing two different, yet related, notions of representing one Hamiltonian with another. The first notion is *encoding*. This is when a Hamiltonian of a certain form exactly represents one of a different form; for example, representing a matrix with complex entries by one with real entries by expressing each complex number as a pair of real numbers. The second notion is *simulation*. This is when the low-energy part of some Hamiltonian is close to an encoding of another Hamiltonian.

After drawing some consequences from our definitions, we go on to introduce our key technique: the use of perturbative gadgets for Hamiltonian simulation. We then show that a number of apparently simple Hamiltonians are universal simulators: they can simulate any other Hamiltonian. Finally, we discuss some corollaries of our results for quantum complexity theory and adiabatic quantum computation.

2 Notation and terminology

As usual, $\mathcal{B}(\mathcal{H})$ denotes the set of linear operators acting on a Hilbert space $\mathcal{H}$. For conciseness, we sometimes also use the notation $\mathcal{M}_n$ for the set of all $n \times n$ matrices with complex entries. Herm_n denotes the subset of all $n \times n$ Hermitian matrices. $\mathbb{1}$ denotes the identity matrix. For integer n, $[n]$ denotes the set $\{1, \ldots, n\}$.

If R, R' are rings, a *ring homomorphism* $\phi : R \to R'$ is a map that is both additive and multiplicative: $\forall a, b \in R : \phi(ab) = \phi(a)\phi(b)$ and $\phi(a + b) = \phi(a) + \phi(b)$. Similarly, a *ring anti-homomorphism* is an additive map that is anti-multiplicative: $\phi(ab) = \phi(b)\phi(a)$. If $\phi(\mathbb{1}) = \mathbb{1}$, we say the map is *unital*.

For a ring R, the corresponding *Jordan ring* R_j is the ring obtained from R by replacing multiplication with Jordan multiplication $\{ab\} := ab + ba$. A *Jordan homomorphism* ϕ on R is an additive map such that $\forall a, b \in R : \phi(ab + ba) = \phi(a)\phi(b) + \phi(b)\phi(a)$. If R is not of characteristic 2, this is equivalent to the constraint that $\forall a \in R : \phi(a^2) = \phi(a)^2$. Note that any ring homomorphism is a Jordan homomorphism, but the converse is not necessarily true.

$\mathrm{spec}(A)$ denotes the spectrum of $A \in \mathcal{M}_n$, i.e. the set of values $\lambda \in \mathbb{C}$ such that $A - \lambda\mathbb{1}$ is not invertible. (This of course coincides with the set of eigenvalues, ignoring multiplicities.) We say that $\phi : \mathcal{M}_n \to \mathcal{M}_m$ is *invertibility-preserving* if $\phi(A)$ is invertible in $\mathcal{M}_m$ for all invertible $A \in \mathcal{M}_n$. We say that ϕ is *spectrum-preserving* if $\mathrm{spec}(\phi(A)) = \mathrm{spec}(A)$ for all $A \in \mathcal{M}_n$.

For an arbitrary Hamiltonian $H \in \mathcal{B}(\mathbb{C}^d)$, we let $P_{\leq \Delta(H)}$ denote the orthogonal projector onto the subspace $S_{\leq \Delta(H)} := \mathrm{span}\{|\psi\rangle : H|\psi\rangle = \lambda|\psi\rangle, \lambda \leq \Delta\}$. We also let $H'|_{\leq \Delta(H)}$ denote the restriction of some other arbitrary Hamiltonian H' to $S_{\leq \Delta(H)}$, and write $H|_{\leq \Delta} := H|_{\leq \Delta(H)}$ and $H_{\leq \Delta} := HP_{\leq \Delta(H)}$.

We say that a Hamiltonian $H \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$ is *k-local* if it can be written as a sum of terms such that each h_i acts non-trivially on at most k subsystems of $(\mathbb{C}^d)^{\otimes n}$. That is, $h_i \in \mathcal{B}((\mathbb{C}^d)^{\otimes k})$ and $H = \sum_i h_i \otimes \mathbb{1}$ where the identity in each term in the sum acts on the subsystems where that h_i does not. An operator on a composite Hilbert space "acts trivially" on the subsystems where

it acts as identity, and "acts non-trivially" on the remaining subsystems. We will often employ a standard abuse of notation, and implicitly exend operators on subsystems to the full Hilbert without explicitly writing the tensor product with identity, allowing us e.g. to write simply $H = \sum h_i$. We say that H is *local* if it is k-local for some k that does not depend on n[1].

We let X, Y, Z denote the Pauli matrices and often follow the condensed-matter convention of writing XX for $X \otimes X$ etc. For example, $XX + YY + ZZ$ is short for $X \otimes X + Y \otimes Y + Z \otimes Z$ and is known as the Heisenberg (exchange) interaction. The XY interaction is $XX + YY$.

Let M be a k-qudit Hermitian matrix. We say that $U \in SU(d)$ *locally diagonalises* M if $U^{\otimes k} M (U^\dagger)^{\otimes k}$ is diagonal. We say that a set $\mathcal{S}$ of Hermitian matrices is simultaneously locally diagonalisable if there exists $U \in SU(d)$ such that U locally diagonalises M for all $M \in \mathcal{S}$. Note that matrices in $\mathcal{S}$ may act on different numbers of qudits, so can be of different sizes.

We will often be interested in families of Hamiltonians. For a subset $\mathcal{S}$ of interactions (Hermitian matrices on a fixed number of qudits), we define the family of $\mathcal{S}$-Hamiltonians to be the set of Hamiltonians which can be written as a sum of interaction terms where each term is either picked from $\mathcal{S}$, with an arbitrary positive or negative real weight, or is an arbitrarily weighted identity term. For example, H is a $\{ZZ\}$-Hamiltonian if it can be written in the form $H = \alpha \mathbb{1} + \sum_{i<j} \beta_{ij} Z_i Z_j$ for some $\alpha, \beta_{ij} \in \mathbb{R}$. A *model* is a (possibly infinite) family of Hamiltonians. Typically the Hamiltonians in a model will be related in some way, e.g. all Hamiltonians with nearest-neighbour Heisenberg interactions on an arbitrarily large 2D lattice (the "2D Heisenberg model").

3 Hamiltonian encodings

Any non-trivial simulation of one Hamiltonian with another will involve encoding the first within the second in some way. Write $H' = \mathcal{E}(H)$ for some "encoding" map $\mathcal{E}$ that encodes a Hamiltonian H into some Hamiltonian H'. Any such encoding should fulfil at least the following basic requirements. First, any observable on the original system should correspond to an observable on the simulator system. Second, the set of possible values of any encoded observable should be the same as for the corresponding original observable. In particular, the energy spectrum of the Hamiltonian should be preserved. Third, the encoding of a probabilistic mixture of observables should be the same as a probabilistic mixture of the encodings of the observables.

To see why this last requirement holds, imagine that we are asked to encode observable A with probability p, and observable B with probability $1 - p$. Then, for any state ρ on the simulator system, the expected value of the encoded observable acting on ρ should be the same as the corresponding probabilistic mixture of the expected values of the encoded observables A and B acting on ρ. In order for this to hold for all states ρ, we need the mixture of observables $pA + (1-p)B$ to be encoded as the corresponding probabilistic mixture of

[1] Technically, this makes sense only for families of Hamiltonians H, where we consider n to be growing.

encodings of A and B.

These operational requirements correspond to the following mathematical requirements on the encoding map $\mathcal{E}$:

1. $\mathcal{E}(A) = \mathcal{E}(A)^\dagger$ for all $A \in \mathrm{Herm}_n$.
2. $\mathrm{spec}(\mathcal{E}(A)) = \mathrm{spec}(A)$ for all $A \in \mathrm{Herm}_n$.
3. $\mathcal{E}(pA + (1-p)B) = p\mathcal{E}(A) + (1-p)\mathcal{E}(B)$ for all $A, B \in \mathrm{Herm}_n$ and all $p \in [0,1]$.

Of course, there are many other desiderata that we would like $\mathcal{E}$ to satisfy, such as preserving the partition function, measurement outcomes, time-evolution, local errors, and others. For the Hamiltonian itself, we almost certainly want $\mathcal{E}$ to not only be convex, but also real-linear: $\mathcal{E}(\sum_i \alpha_i h_i) = \sum_i \alpha_i \mathcal{E}(h_i)$, so that a Hamiltonian expressed as a sum of terms can be encoded by encoding the terms separately. However, we will see later that meeting just the above three basic requirements necessarily implies also meeting all these other operational requirements (which we will make precise).

It turns out there is a simple and elegant characterisation of what such encodings have to look like. To prove this, we will need the following theorem concerning Jordan ring homomorphisms.

Theorem 1 (follows from [JR52], Theorem 4 and [Mar67], Theorem 2) *For any $n \geq 2$, any Jordan homomorphism of the Jordan ring Herm_n can be extended in one and only one way to a homomorphism of the matrix ring $\mathcal{M}_n$.*

Theorem 1 was shown by Jacobson and Rickart for $n \geq 3$ [JR52], and by Martindale for $n = 2$ [Mar67], in each case in a far more general setting than we need here.

Lemma 2 *Any unital, invertibility-preserving, real-linear map $\phi : \mathrm{Herm}_n \to \mathrm{Herm}_m$ is a Jordan homomorphism.*

Proof The argument is standard (see e.g. [HŠ03]).

$\phi(H - \lambda\mathbb{1}) = \phi(H) - \lambda\mathbb{1}$, thus $\mathrm{spec}(\phi(H)) \subseteq \mathrm{spec}(H)$ since ϕ is invertibility-preserving. In particular, $\mathrm{spec}(\phi(P)) \in \{0,1\}$ for every projector P. Since $\phi(P)$ is also Hermitian, this implies $\phi(P)$ is a projector.

By the spectral decomposition, any $H \in \mathrm{Herm}_n$ can be decomposed as $H = \sum_i \lambda_i P_i$ where P_i are mutually orthogonal projectors and $\lambda_i \in \mathbb{R}$. For $i \neq j$, $P_i + P_j$ is a projector, thus $\phi(P_i + P_j)$ is a projector and $(\phi(P_i + P_j))^2 = \phi(P_i) + \phi(P_j)$, so that $\phi(P_i)\phi(P_j) + \phi(P_i)\phi(P_j) = 0$. Therefore, $\phi(H)^2 = \sum_i \lambda_i^2 \phi(P_i)^2 + \sum_{i \neq j} \lambda_i \lambda_j \phi(P_i)\phi(P_j) = \sum_i \lambda_i^2 \phi(P_i) = \phi(H^2)$. □

Theorem 3 (Encodings) *For any map $\mathcal{E} : \mathrm{Herm}_n \to \mathrm{Herm}_m$, the following are equivalent:*

(i). For all $A, B \in \mathrm{Herm}_n$, and all $p \in [0,1]$:

1. $\mathcal{E}(A) = \mathcal{E}(A)^\dagger$

2. $\mathrm{spec}(\mathcal{E}(A)) = \mathrm{spec}(A)$
3. $\mathcal{E}(pA + (1-p)B) = p\mathcal{E}(A) + (1-p)\mathcal{E}(B)$.

(ii). *There exists a unique extension $\mathcal{E}' : \mathcal{M}_n \to \mathcal{M}_m$ such that $\mathcal{E}'(H) = \mathcal{E}(H)$ for all $H \in \mathrm{Herm}_n$ and, for all $A, B \in \mathcal{M}_n$ and $x \in \mathbb{R}$:*

a. $\mathcal{E}'(\mathbb{1}) = \mathbb{1}$
b. $\mathcal{E}'(A^\dagger) = \mathcal{E}'(A)^\dagger$
c. $\mathcal{E}'(A + B) = \mathcal{E}'(A) + \mathcal{E}'(B)$
d. $\mathcal{E}'(AB) = \mathcal{E}'(A)\mathcal{E}'(B)$
e. $\mathcal{E}'(xA) = x\mathcal{E}'(A)$.

(iii). *There exists a unique extension $\mathcal{E}' : \mathcal{M}_n \to \mathcal{M}_m$ such that $\mathcal{E}'(H) = \mathcal{E}(H)$ for all $H \in \mathrm{Herm}_n$ with $\mathcal{E}'$ of the form*

$$\mathcal{E}'(M) = U\left(M^{\oplus p} \oplus \bar{M}^{\oplus q}\right)U^\dagger \tag{1}$$

for some non-negative integers p, q and unitary $U \in \mathcal{M}_m$, where $M^{\oplus p} := \bigoplus_{i=1}^{p} M$ and $\bar{M}$ denotes complex conjugation.

We call a map $\mathcal{E}$ satisfying (i) to (iii) an encoding.

Note that (iii) is basis-independent, despite the occurrence of complex conjugation; taking the complex conjugation with respect to a different basis is equivalent to modifying U, which just gives another encoding. Given that $\mathcal{E}'$ is unique, for the remainder of the paper we simply identify $\mathcal{E}'$ with $\mathcal{E}$. In particular, this allows us to assume that $\mathcal{E}$ is of the form specified in part (iii). The characterisation (1) can equivalently be written as

$$\mathcal{E}'(M) = U\left(M \otimes P + \bar{M} \otimes Q\right)U^\dagger \tag{2}$$

for some orthogonal projectors P and Q such that $P + Q = \mathbb{1}$; this alternative form will sometimes be useful below. We think of the system on which P and Q act as an ancilla, and often label this "extra" subsystem by the letter E.

Proof (i) $\Rightarrow$ (ii):
We first show that $\mathcal{E}$ is a Jordan homomorphism. Condition (i)1 states that $\mathcal{E}$ preserves Herm_n, and condition (i)2 implies that $\mathcal{E}$ is unital and invertibility-preserving on Herm_n, with $\mathcal{E}(0) = 0$. We next check that $\mathcal{E}(0) = 0$ together with condition (i)3 are equivalent to real-linearity of $\mathcal{E}$. For any $\lambda < 0$, setting $p = \lambda/(\lambda - 1)$, $B = pA/(p-1)$ and using condition (i)3 gives

$$0 = \mathcal{E}(0) = p\mathcal{E}(A) + (1-p)\mathcal{E}(pA/(p-1)) \;\Leftrightarrow\; \lambda\mathcal{E}(A) = \mathcal{E}(\lambda A). \tag{3}$$

Apply (3) to λA to get $\mathcal{E}(\lambda^2 A) = \lambda^2\mathcal{E}(A)$, showing that $\mathcal{E}$ is homogeneous for all real scalars. Additivity follows by combining condition (i)3 and homogeneity: $\mathcal{E}(A + B) = \mathcal{E}(2A)/2 + \mathcal{E}(2B)/2 = \mathcal{E}(A) + \mathcal{E}(B)$. Therefore $\mathcal{E}$ is also real-linear so by Lemma 2 $\mathcal{E}$ is a Jordan homomorphism.

By Theorem 1, there exists a unique homomorphism $\mathcal{E}' : \mathcal{M}_n \to \mathcal{M}_m$ such that $\mathcal{E}'(H) = \mathcal{E}(H)$ for all $H \in \mathrm{Herm}_n$. As $\mathcal{E}'$ agrees with $\mathcal{E}$ on Herm_n, it satisfies (ii)a. As $\mathcal{E}'$ is a homomorphism, it satisfies (ii)c and (ii)d by definition; this also implies that $\mathcal{E}'(xA) = \mathcal{E}'(x\mathbb{1})\mathcal{E}'(A) = \mathcal{E}(x\mathbb{1})\mathcal{E}'(A) = x\mathcal{E}'(A)$ for any $x \in \mathbb{R}$, so (ii)e holds.

We finally prove (ii)b. It is sufficient to show that $\mathcal{E}'(i\mathbb{1})^\dagger = -\mathcal{E}'(i\mathbb{1})$, because if this holds we can expand any matrix $A \in \mathcal{M}_n$ as $A = B + iC$ for some Hermitian matrices B and C to obtain

$$\mathcal{E}'(A^\dagger) = \mathcal{E}'(B - iC) = \mathcal{E}'(B) - \mathcal{E}'(C)\mathcal{E}'(i\mathbb{1}) = \mathcal{E}'(B)^\dagger + \mathcal{E}'(C)^\dagger\mathcal{E}'(i\mathbb{1})^\dagger \tag{4}$$

$$= \mathcal{E}'(B + iC)^\dagger = \mathcal{E}'(A)^\dagger. \tag{5}$$

To show $\mathcal{E}'(i\mathbb{1})^\dagger = -\mathcal{E}'(i\mathbb{1})$, we first write $i\mathbb{1}$ as a linear combination of products of Hermitian matrices. That this can be done is an immediate consequence of the fact that $\mathcal{M}_n$ is the enveloping associative ring of Herm_n. However, it can also be seen explicitly by writing

$$i\,|j\rangle\langle j| = |j\rangle\langle j|\,(i\,|j\rangle\langle k| - i\,|k\rangle\langle j|)(|j\rangle\langle k| + |k\rangle\langle j|) \tag{6}$$

for any j, and some $k \neq j$; summing this product over j, we obtain $i\mathbb{1}$. Thus we can write $i\mathbb{1} = \sum_j A_j B_j C_j$ for Hermitian matrices A_j, B_j, C_j. By taking adjoints on both sides, it follows that $-i\mathbb{1} = \sum_j C_j B_j A_j$. So we have

$$\mathcal{E}'(i\mathbb{1})^\dagger = \mathcal{E}'\Big(\sum_j A_j B_j C_j\Big)^\dagger = \Big(\sum_j \mathcal{E}(A_j)\mathcal{E}(B_j)\mathcal{E}(C_j)\Big)^\dagger \tag{7}$$

$$= \sum_j \mathcal{E}(C_j)\mathcal{E}(B_j)\mathcal{E}(A_j) = \mathcal{E}'\Big(\sum_j C_j B_j A_j\Big) \tag{8}$$

$$= \mathcal{E}'(-i\mathbb{1}) = -\mathcal{E}'(i\mathbb{1}). \tag{9}$$

(ii) $\Rightarrow$ (iii):
Existence and uniqueness of $\mathcal{E}'$ were already shown in the previous part. In the proof of the remaining claim, for readability we just use $\mathcal{E}$ to denote this unique extension. First define the complex structure $J := \mathcal{E}(i\mathbb{1}) \equiv \mathcal{E}(i)$ (where the latter notation is a convenient shorthand). We have

$$J^2 = \mathcal{E}(i)\mathcal{E}(i) = \mathcal{E}(i^2) = \mathcal{E}(-1) = -\mathbb{1}, \tag{10}$$

thus J has eigenvalues $\pm i$. Furthermore,

$$J^\dagger = \mathcal{E}(i)^\dagger = \mathcal{E}(i^\dagger) = -\mathcal{E}(i) = -J, \tag{11}$$

so J is anti-Hermitian, hence diagonalisable by a unitary transformation.

For any $A \in \mathrm{Herm}_n$, we have

$$J\mathcal{E}(A) = \mathcal{E}(i)\mathcal{E}(A) = \mathcal{E}(iA) = \mathcal{E}(Ai) = \mathcal{E}(A)J, \tag{12}$$

so that $[\mathcal{E}(A), J] = 0$. Thus $\mathcal{E}(A)$ and J are simultaneously diagonalisable for all A. $\mathcal{H} = \mathcal{H}_+ \oplus \mathcal{H}_-$ therefore decomposes into a direct sum of the $\pm i$ eigenspaces of J, on which $\mathcal{E}(A) = A_+ \oplus A_-$ acts invariantly.

Now, restricting to either of these invariant subspaces,

$$\mathcal{E}(A)|_\pm = A_\pm \tag{13}$$
$$\mathcal{E}(iA)|_\pm = JA_\pm = \pm iA_\pm \tag{14}$$
$$\mathcal{E}(AB)|_\pm = \mathcal{E}(A)\mathcal{E}(B)|_\pm = A_\pm B_\pm \tag{15}$$
$$\mathcal{E}(A^\dagger)|_\pm = \mathcal{E}(A)^\dagger|_\pm = A_\pm^\dagger. \tag{16}$$

Thus $\mathcal{E} = \mathcal{E}_+ \oplus \mathcal{E}_-$ decomposes into a direct sum of a *-representation $\mathcal{E}_+(A) := \mathcal{E}(A)|_+$ and an anti-*-representation[1] $\mathcal{E}_-(A) := \mathcal{E}(A)|_-$. Since for any vector $|\psi\rangle \in \mathbb{C}^m$, $\mathcal{E}_\pm(\mathbb{1})|\psi\rangle = \mathbb{1}|\psi\rangle = |\psi\rangle$, these (anti-)*-representations are necessarily non-degenerate.

By a standard result on the representations of finite-dimensional C*-algebras [Dav91, Corollary III.1.2], any non-degenerate *-representation of $\mathcal{M}_n$ is unitarily equivalent to a direct sum of identity representations. If ϕ is an anti-*-homomorphism, let $\varphi(A) := \overline{\phi(A)}$. Then $\varphi(iA) = \overline{\phi(iA)} = \overline{-i\phi(A)} = i\varphi(A)$, $\varphi(A + B) = \varphi(A) + \varphi(B)$, $\varphi(A^\dagger) = \varphi(A)^\dagger$, and $\varphi(AB) = \varphi(A)\varphi(B)$. Thus $\phi(A) = \overline{\varphi(A)}$ where φ is a *-homomorphism. Therefore, any non-degenerate anti-*-representation is unitarily equivalent to a direct sum of complex conjugates of identity representations, which completes the argument.

(iii) ⇒ (i) can readily be verified directly. □

The above theorem characterises encodings of observables. This immediately tells us how to encode physical systems themselves, expressed as Hamiltonians: since the Hamiltonian itself is an observable, the encoding map must have the same characterisation.

It is easy to see from the characterisation in part (iii) of the Theorem that any encoding preserves all interesting physical properties of the original Hamiltonian. For example, the set of eigenvalues is preserved, up to possibly duplicating each eigenvalue the same number of times, implying preservation of the partition function (up to an unimportant constant factor). It is also easy to see that any encoding $\mathcal{E}$ properly encodes arbitrary quantum channels: if $\{E_k : \sum_k E_k^\dagger E_k = \mathbb{1}\}$ are the Kraus operators of the channel, then

$$\sum_k \mathcal{E}(E_k)^\dagger \mathcal{E}(E_k) = \mathbb{1}. \tag{17}$$

3.1 A map on states, $\mathcal{E}_{\text{state}}$

We now show that, for any encoding $\mathcal{E}$, there exists a corresponding map $\mathcal{E}_{\text{state}}$ that encodes quantum states ρ such that encoded observables $\mathcal{E}(A)$ applied to encoded states $\mathcal{E}_{\text{state}}(\rho)$ have correct expectation values.

[1] By "anti-*-representation" we mean an anti-linear algebra homomorphism, *not* a *-antihomomorphism (which would be a linear map preserving adjoints that reverses the order of multiplication).

First, note that for any observable A and any state ρ' on the simulator system, we have

$$\begin{aligned}\mathrm{Tr}(\mathcal{E}(A)\rho') &= \mathrm{Tr}[U(A\otimes P + \bar{A}\otimes Q)U^\dagger\rho'] &(18)\\ &= \mathrm{Tr}[(A\otimes\mathbb{1})(\mathbb{1}\otimes P)U^\dagger\rho' U)] + \mathrm{Tr}[(\bar{A}\otimes\mathbb{1})(\mathbb{1}\otimes Q)U^\dagger\rho' U] &(19)\\ &= \mathrm{Tr}[AF(\rho')] + \mathrm{Tr}[\bar{A}\,\overline{B(\rho')}] = \mathrm{Tr}(A\rho) &(20)\end{aligned}$$

where

$$F(\rho') = \mathrm{Tr}_E[(\mathbb{1}\otimes P)U^\dagger\rho' U], \qquad B(\rho') = \overline{\mathrm{Tr}_E[(\mathbb{1}\otimes Q)U^\dagger\rho' U]}, \tag{21}$$

$$\rho = F(\rho') + B(\rho') \tag{22}$$

and we label the second subsystem E as discussed after (2). Note that $F(\rho')$ and $B(\rho')$ are both positive but not necessarily normalised, but ρ is normalised.

Therefore any map $\mathcal{E}_{\mathrm{state}}(\rho)$ on states ρ such that $\rho = F(\mathcal{E}_{\mathrm{state}}(\rho))+B(\mathcal{E}_{\mathrm{state}}(\rho))$ will preserve measurement outcomes appropriately. One natural choice is

$$\mathcal{E}_{\mathrm{state}}(\rho) = \begin{cases} U(\rho\otimes\sigma)U^\dagger \text{ for some } \sigma \text{ such that } P\sigma=\sigma & \text{if } P\neq 0\\ U(\bar{\rho}\otimes\sigma)U^\dagger \text{ for some } \sigma \text{ such that } Q\sigma=\sigma & \text{otherwise.}\end{cases} \tag{23}$$

Then in the former case $F(\mathcal{E}_{\mathrm{state}}(\rho)) = \rho$, $B(\mathcal{E}_{\mathrm{state}}(\rho)) = 0$; and in the latter case the roles of F and B are reversed.

We now show that $\mathcal{E}_{\mathrm{state}}$ simulates time-evolution correctly too. We have

$$F(e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}) = e^{-iHt}F(\rho')e^{iHt}, \tag{24}$$

$$B(e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}) = e^{iHt}B(\rho')e^{-iHt}. \tag{25}$$

This is why they are labelled with the letters F and B: the F part evolves *forwards* in time while the B part evolves *backwards* in time. Taking $\rho' = \mathcal{E}_{\mathrm{state}}(\rho)$, we have proven the following result.

Proposition 4 *For any encoding $\mathcal{E}$, the corresponding map $\mathcal{E}_{\mathrm{state}}$ satisfies the following:*

(i). $\mathrm{Tr}\left(\mathcal{E}(A)\mathcal{E}_{\mathrm{state}}(\rho)\right) = \mathrm{Tr}(A\rho)$

(ii). For any time t,

$$e^{-i\mathcal{E}(H)t}\mathcal{E}_{\mathrm{state}}(\rho)e^{i\mathcal{E}(H)t} = \begin{cases}\mathcal{E}_{\mathrm{state}}(e^{-iHt}\rho e^{iHt}) & \text{if } p\geq 1\\ \mathcal{E}_{\mathrm{state}}(e^{iHt}\rho e^{-iHt}) & \text{if } p = 0.\end{cases} \tag{26}$$

It is worth highlighting the last point. We see that if $p \geq 1$, evolving according to $\mathcal{E}(H)$ for time t simulates evolving according to H for time t, as we would expect; but that if $p = 0$, we simulate evolution according to H for time $-t$. That is, if our encoding only includes copies of $\bar{H}$, we simulate evolution backwards in time. To avoid this issue, we define the concept of a *standard* encoding as one where $p \geq 1$, and hence which is able to simulate evolution forward in time.

Definition 5 (Standard encoding) *An encoding $\mathcal{E}(M) = U(M^{\oplus p}\oplus\bar{M}^{\oplus q})U^\dagger$ is a* standard encoding *if $p \geq 1$.*

3.1.1 Gibbs-preserving state mappings

The choice of $\mathcal{E}_{\text{state}}$ in (23) is convenient, as it allows us to use the same mapping $\mathcal{E}$ for both the Hamiltonian and for observables. However, it does not map Gibbs states $e^{-\beta H}/\operatorname{Tr}(e^{-\beta H})$ of the original system to Gibbs states $e^{-\beta' H'}/\operatorname{Tr}(e^{-\beta' H'})$ of the simulator. If we have limited ability to manipulate or prepare states of the simulator, it may be difficult to prepare a state of the form (23). At equilibrium, the system will naturally be in a Gibbs state. From this perspective, it would be more natural if the state mapping identified Gibbs states of the original system with Gibbs states of the simulator.

An alternative choice of $\mathcal{E}_{\text{state}}$ *does* map Gibbs states to Gibbs states:

$$\mathcal{E}_{\text{state}}(\rho) = \frac{\mathcal{E}(\rho)}{\operatorname{Tr}[\mathcal{E}(\rho)]} = \frac{1}{p+q} U(\rho \otimes P + \bar{\rho} \otimes Q)U^\dagger \tag{27}$$

where $p = \operatorname{Tr}(P)$ and $q = \operatorname{Tr}(Q)$. However, to obtain the correct measurement outcome probabilities, we now need to choose a slightly different mapping for observables:[1]

$$\mathcal{E}_{\text{meas}}(A) = \begin{cases} \frac{p+q}{p} U(A \otimes P)U^\dagger \text{ if } P \neq 0 \\ \frac{p+q}{q} U(\bar{A} \otimes Q)U^\dagger \text{ otherwise.} \end{cases} \tag{28}$$

For simplicity, in the remainder of the paper we will state and prove our results for the choice of state mapping $\mathcal{E}_{\text{state}}$ from (23), so that both Hamiltonians and observables are encoded by $\mathcal{E}$. However, our results also go through with the appropriate minor modifications for the choice of Gibbs-preserving $\mathcal{E}_{\text{state}}$ from (27), where the simulator Hamiltonian is still constructed using $\mathcal{E}$ but observables are encoded by the $\mathcal{E}_{\text{meas}}$ from (28).

Note that $\mathcal{E}_{\text{meas}}$ has been chosen so that measuring $\mathcal{E}_{\text{meas}}(A)$ will only pick up the $F(\rho')$ part of a state ρ' on the simulator. We therefore include results concerning the behaviour of F, in order to cover the choice of $\mathcal{E}_{\text{state}}$ given in (27), as well other mappings on states.

3.2 The complex-to-real encoding

The only nontrivial encoding (as opposed to simulation, q.v.) that we will need to use is an encoding of complex Hamiltonians as real Hamiltonians.

Lemma 6 *There exists an encoding φ such that for any Hamiltonian $H \in \mathcal{B}(\mathbb{C}^d)$, the encoded Hamiltonian $H' = \varphi(H) \in \mathcal{B}(\mathbb{R}^{2d})$ is* real.

Proof This follows from the canonical Hilbert space isomorphism $\mathbb{C}^d \simeq \mathbb{R}^{2d}$ where the latter is endowed with a linear complex structure J.

Concretely, let

$$J := \begin{pmatrix} 0 & \mathbb{1}_d \\ -\mathbb{1}_d & 0 \end{pmatrix} = iY \otimes \mathbb{1}_d \tag{29}$$

[1] The Hamiltonian is of course also an observable. With this choice of state mapping, to construct the simulator Hamiltonian we must still use the mapping $H' = \mathcal{E}(H)$. But if we want to *measure* the Hamiltonian – i.e. carry out the measurement on the simulator that corresponds to measuring the energy of the original system – we must measure $\mathcal{E}_{\text{meas}}(H)$.

where where $\mathbb{1}_d$ is the $d \times d$ identity matrix, and define the mapping

$$\begin{array}{rccl} \varphi: & \mathcal{B}(\mathbb{C}^d) & \to & \mathcal{B}(\mathbb{R}^{2d}) \\ & \varphi(M) & = & \operatorname{Re} M \oplus \operatorname{Re} M + J \operatorname{Im} M \oplus \operatorname{Im} M. \end{array} \tag{30}$$

To see that φ is indeed a valid encoding, we can either verify directly that it satisfies all the properties listed in part (i) of Theorem 3, or observe that

$$\varphi(M) = U(M \oplus \bar{M})U^\dagger \quad \text{where} \quad U = \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{1} & \mathbb{1} \\ i\mathbb{1} & -i\mathbb{1} \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \otimes \mathbb{1}, \tag{31}$$

which is manifestly of the form given in part (iii) of Theorem 3. The Lemma follows by setting

$$H' = \varphi(H) = \operatorname{Re}(H) \oplus \operatorname{Re}(H) + J \operatorname{Im}(H) \oplus \operatorname{Im}(H). \tag{32}$$

□

When applied to a Hamiltonian on a system of n qubits, the encoding of Lemma 6 is local (see Section 3.3). Indeed, it produces a Hamiltonian H' on $n+1$ qubits, given by

$$H' = |+_y\rangle\langle+_y| \otimes H + |-_y\rangle\langle-_y| \otimes \bar{H} \tag{33}$$

where $|\pm_y\rangle = (|0\rangle \pm i\,|1\rangle)/\sqrt{2}$ are the eigenstates of Y. It is easy to see that H' is real since $\overline{|+_y\rangle} = |-_y\rangle$. Any complex k-local interaction is mapped to a $(k+1)$-local interaction involving the additional qubit.

This additional qubit therefore has a special significance in the construction, which leads to two unwanted consequences. Firstly, the interaction graph of H' is in general more complicated than that of H. Any geometric locality or spatial sparsity in the original Hamiltonian H is lost, as all complex local terms are mapped to interactions in H' that involve this additional qubit. Secondly, an error on this single additional qubit would mix the spaces where H and $\bar{H}$ act. This could lead to unusual errors when simulating the time evolution of ρ under H with the simulator H'.

In Lemma 21 below we give an alternative to this encoding that avoids these problems.

3.3 Local encodings

So far, we have considered encodings of arbitrary Hamiltonians, with no additional structure. However, in Hamiltonian simulation, we are typically interested in many-body Hamiltonians composed of local interactions between subsets of particles. That is, Hamiltonians $H \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$ with $H = \sum_i h_{S_i}$, where the local terms $h_{S_i} \in \mathcal{B}((\mathbb{C}^d)^{\otimes |S_i|})$ act on subsets S_i of the particles (implicitly extended to $\mathcal{B}((\mathbb{C}^d)^{\otimes n})$ in the sum by tensoring with identity on the rest of the space, as usual).

In this case, we typically want our encoding to be *local*, i.e. it should map local observables to local observables, and consequently

$$\mathcal{E}(h_{S_i} \otimes \mathbb{1}) = h'_{S'_i} \otimes \mathbb{1} \tag{34}$$

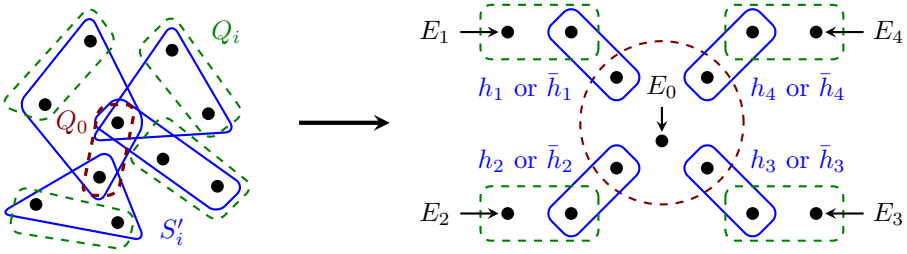


Figure 1: Any local encoding can be decoupled into disjoint subsystems by local unitaries on the Q_i systems. Each subsystem encodes one of the qudits of the original system. Here S'_i denotes the subsystems encoding qudit i as a direct sum of identity and conjugate representations. E_i denotes ancilla subsystems.

so that the simulation $H' = \mathcal{E}(H) = \sum_i h'_{S'_i}$ is itself a local Hamiltonian.

Definition 7 (Local encoding) *Let $\mathcal{E} : \mathcal{B}(\bigotimes_{i=1}^{n} \mathcal{H}_i) \to \mathcal{B}(\bigotimes_{i=1}^{n'} \mathcal{H}'_i)$ be an encoding, and let $\{S'_i\}_{i=1}^{n}$ be subsets of $[n']$. We say that the encoding is* local *with respect to $\{S'_i\}$ if for any operator $A \in \mathcal{B}(\mathcal{H}_i)$, $\mathcal{E}(A \otimes \mathbb{1})$ acts non-trivially only on S'_i.*

Theorem 8 *Let $\mathcal{E} : \mathcal{B}(\bigotimes_{i=1}^{n} \mathcal{H}_i) \to \mathcal{B}(\bigotimes_{i=1}^{n'} \mathcal{H}'_i)$ be a local encoding with respect to $\{S'_i\}$. Denote $Q_0 = \bigcup_{i,j} S'_i \cap S'_j$ and $Q_i = S'_i \setminus Q_0$ (see Figure 1). Then there exist decompositions $\mathcal{H}_{Q_0} \simeq E_0 \otimes (\bigotimes_i \mathcal{H}_i^{(\mathrm{in})})$ and $\mathcal{H}_{Q_i} \simeq E_i \otimes \mathcal{H}_i^{(\mathrm{out})}$, together with identifications $\mathcal{H}_i \simeq \mathcal{H}_i^{(\mathrm{in})} \otimes \mathcal{H}_i^{(\mathrm{out})}$ and a decomposition $E_0 = \bigoplus_\alpha \left(\bigotimes_{i=0}^{n} E_{0.i}^{(\alpha)} \right)$, such that the encoding takes the form*

$$\mathcal{E}(M) = U_{Q_0} \bigoplus_\alpha \left(\left[\bigotimes_i U_{(i)}^{(\alpha)} \right] \left(M^{(\alpha)} \otimes \mathbb{1}_{E_1,\ldots,E_n} \otimes \mathbb{1}_{E_{0.1}^{(\alpha)},\ldots,E_{0.n}^{(\alpha)}} \right) \left[\bigotimes_i U_{(i)}^{(\alpha)} \right]^\dagger \right) U_{Q_0}^\dagger \tag{35}$$

where U_{Q_0} acts non-trivially only on $\mathcal{H}_{Q_0}$, each $U_{Q_i^+}^{(\alpha)}$ acts on $\mathcal{H}_{Q_i^+}^{(\alpha)} := \mathcal{H}_i \otimes E_i \otimes E_{0.i}^{(\alpha)}$, and each $M^{(\alpha)} = M$ or $\bar{M}$.

Theorem 8 implies that the locality structure of an encoding is fully determined by how it maps 1-local operators. Note that any of the Hilbert spaces in the decomposition could be one-dimensional.

The characterisation in Theorem 8 shows that the most general possible encoding of local Hamiltonians looks very like the complex-to-real encoding from Lemma 6. Up to local unitaries, local encodings are just direct sums of product encodings, with a classical ancilla that determines whether to take the complex conjugate of all the local interactions or not.

To prove Theorem 8, we will need the following (slightly generalised) lemma from [AE11], which is itself a special case of a result from [BV05]:

Lemma 9 (Lemma 3.3 of [AE11]) *Let $\mathcal{H} = \bigotimes_{i=0}^{n} \mathcal{H}_i$ be a Hilbert space and let $\mathcal{A}_{0,k}$, $k \in \{1, \ldots, n\}$, be sets of matrices which act non-trivially only on $\mathcal{H}_0$ and $\mathcal{H}_k$, such that matrices from different sets all commute. Then there exists a direct sum decomposition of $\mathcal{H}_0$*

$$\mathcal{H}_0 = \bigoplus_{\alpha} \mathcal{H}_0^{(\alpha)} \tag{36}$$

such that inside each subspace $\mathcal{H}_0^{(\alpha)}$ there is a tensor product structure

$$\mathcal{H}_0^{(\alpha)} = \bigotimes_{i=0}^{n} \mathcal{H}_{0.i}^{(\alpha)}, \tag{37}$$

and any element $A \in \mathcal{A}_{0,k}$ preserves the subspaces $\mathcal{H}^{(\alpha)} := \mathcal{H}_0^{(\alpha)} \otimes \bigotimes_{i=1}^{n} \mathcal{H}_i$. Moreover $A|_{\mathcal{H}^{(\alpha)}}$ acts non-trivially only on $\mathcal{H}_{0.k}^{(\alpha)} \otimes \mathcal{H}_k$.

In [AE11] this lemma is stated only in terms of single operators $H_{0,k}$ rather than sets of operators $\mathcal{A}_{0,k}$, but the proof from [AE11] or [BV05] can be easily seen to generalise to this case.

Proof (of Theorem 8) Let $\mathcal{A}_i = \langle \mathcal{E}(A_i \otimes \mathbb{1}) : A \in \mathcal{B}(\mathcal{H}_i) \rangle$ be the algebra generated by the operators $\{\mathcal{E}(A_i \otimes \mathbb{1})\}$. By assumption, $\mathcal{A}_i$ acts non-trivially only on $\mathcal{H}_{Q_0 \cup Q_i}$. Multiplicativity of encodings (Theorem 3(ii)d) yields that, for $i \neq j$ and all $A \in \mathcal{B}(\mathcal{H}_i)$, $B \in \mathcal{B}(\mathcal{H}_j)$,

$$[\mathcal{E}(A_i \otimes \mathbb{1}), \mathcal{E}(B_j \otimes \mathbb{1})] = \mathcal{E}([A_i \otimes \mathbb{1}, B_j \otimes \mathbb{1}]) = 0. \tag{38}$$

Thus the algebras $\mathcal{A}_i$ fulfil the hypothesis of Lemma 9 for the Hilbert spaces $\mathcal{H} = \bigotimes_{i=1}^{n} \mathcal{H}_{Q_i}$. Applying Lemma 9, we obtain a decomposition

$$\mathcal{H}_{Q_0} = \bigoplus_{\alpha} \left[\bigotimes_{i=0}^{n} \mathcal{H}_{0.i}^{(\alpha)} \right] \tag{39}$$

such that $\mathcal{A}_i = \bigoplus_{\alpha} \mathcal{A}_i^{(\alpha)}$ where $\mathcal{A}_i^{(\alpha)}$ acts non-trivially only on the factors $\mathcal{H}_{Q_i} \otimes \mathcal{H}_{0.i}^{(\alpha)}$. Let $U_{Q_0}^{\dagger} : \mathcal{H}_{Q_0} \to \bigoplus_{\alpha} \bigotimes_{i=0}^{n} \mathcal{H}_{0.i}^{(\alpha)}$ be the unitary change of basis corresponding to this decomposition of $\mathcal{H}_{Q_0}$.

Now, from the general characterisation of encodings (1), we know $\mathcal{A}_i$ has the form

$$\mathcal{A}_i = \left\langle W \left((A \otimes \mathbb{1})^{\oplus p} \oplus (\bar{A} \otimes \mathbb{1})^{\oplus q} \right) W^{\dagger} \right\rangle = \left\langle W (A^{\oplus Dp} \oplus \bar{A}^{\oplus Dq}) W^{\dagger} \right\rangle \tag{40}$$

for some unitary W and $p, q \in \mathbb{N}$. (D here is the dimension of the identity operator which acts on all but the i'th qudit of the original system.) Thus $\mathcal{A}_i$ is unitarily equivalent to a direct sum of identity and conjugated identity

representations of the full matrix algebra on $\mathcal{H}_i$. Note that this decomposes $\mathcal{A}_i$ into irreducible representations, as the full matrix algebra in any dimension is irreducible.

Since $\mathcal{A}_i$ is simultaneously equivalent to $\bigoplus_\alpha \mathcal{A}_i^{(\alpha)}$, each $\mathcal{A}_i^{(\alpha)}$ must itself be unitarily equivalent to a direct sum of copies of identity and conjugated identity representations. Thus, for arbitrary $A \in \mathcal{B}(\mathcal{H}_i)$,

$$\mathcal{E}(A_i \otimes \mathbb{1}) = U_{Q_0} \left(\bigoplus_\alpha \left[U_{Q_i^+}^{(\alpha)} \left(A^{\oplus n_i(\alpha)} \oplus \bar{A}^{\oplus m_i(\alpha)} \right) U_{Q_i^+}^{(\alpha)\dagger} \otimes \mathbb{1}_{\text{rest}}^{(\alpha)} \right] \right) U_{Q_0}^\dagger \quad (41)$$

for some $n_i(\alpha)$, $m_i(\alpha) \in \mathbb{N}$, where $U_{Q_i^+}^{(\alpha)}$ acts on $\mathcal{H}_{Q_i^+}^{(\alpha)} := \mathcal{H}_{Q_i} \otimes \mathcal{H}_{0.i}^{(\alpha)}$.

We will show that for each α, either $n_i(\alpha) = 0$ for all i, or $m_i(\alpha) = 0$ for all i. Note that $J = \mathcal{E}(i\mathbb{1}) = \mathcal{E}\left((i\mathbb{1}_j) \otimes \mathbb{1}_k \otimes \mathbb{1}_{\text{rest}}\right) = \mathcal{E}\left(\mathbb{1}_j \otimes (i\mathbb{1}_k) \otimes \mathbb{1}_{\text{rest}}\right)$ for any qudits j, k of the original system. From (41),

$$\begin{aligned} &\left(\bigoplus_\alpha U_{Q_j^+}^{(\alpha)} \otimes U_{Q_k^+}^{(\alpha)} \right)^\dagger U_{Q_0}^\dagger \, \mathcal{E}((i\mathbb{1}_j) \otimes \mathbb{1}_k \otimes \mathbb{1}_{\text{rest}}) \, U_{Q_0} \left(\bigoplus_\alpha U_{Q_j^+}^{(\alpha)} \otimes U_{Q_k^+}^{(\alpha)} \right) \\ &\qquad = i \bigoplus_\alpha \left[\left(\mathbb{1}^{\oplus n_j(\alpha)} \oplus (-\mathbb{1})^{\oplus m_j(\alpha)} \right) \otimes \mathbb{1}_{Q_k^+}^{(\alpha)} \otimes \mathbb{1}_{\text{rest}}^{(\alpha)} \right]. \quad (42) \end{aligned}$$

Equating this with $U_{Q_0}^\dagger \mathcal{E}(\mathbb{1}_j \otimes (i\mathbb{1}_k) \otimes \mathbb{1}_{\text{rest}}) U_{Q_0}$ and matching up factors in the direct sum over α, we obtain

$$\left(\mathbb{1}^{\oplus n_j(\alpha)} \oplus (-\mathbb{1})^{\oplus m_j(\alpha)} \right) \otimes \mathbb{1}_{Q_k^+}^{(\alpha)} = \mathbb{1}_{Q_j^+}^{(\alpha)} \otimes \left(\mathbb{1}^{\oplus n_k(\alpha)} \oplus (-\mathbb{1})^{\oplus m_k(\alpha)} \right), \quad (43)$$

which is only possible if either $n_j(\alpha) = n_k(\alpha) = 0$ or $m_j(\alpha) = m_k(\alpha) = 0$. Since this holds for any pair j, k, either $n_i(\alpha) = 0$ for all i, or $m_i(\alpha) = 0$ for all i, as claimed. We write "$n(\alpha) = 0$", "$m(\alpha) = 0$" as shorthand for each of these two cases. Then

$$\begin{aligned} U_{Q_0}^\dagger \mathcal{E}(A_i \otimes \mathbb{1}) U_{Q_0} = &\left(\bigoplus_{\alpha : m(\alpha) = 0} U_{Q_i^+}^{(\alpha)} (A \otimes \mathbb{1}_{E_i^+}^{(\alpha)}) U_{Q_i^+}^{(\alpha)\dagger} \otimes \mathbb{1}_{\text{rest}}^{(\alpha)} \right) \\ &\oplus \left(\bigoplus_{\alpha : n(\alpha) = 0} U_{Q_i^+}^{(\alpha)} (\bar{A} \otimes \mathbb{1}_{E_i^+}^{(\alpha)}) U_{Q_i^+}^{(\alpha)\dagger} \otimes \mathbb{1}_{\text{rest}}^{(\alpha)} \right) \quad (44) \end{aligned}$$

where $\dim(\mathbb{1}_{E_i^+}^{(\alpha)}) = n_i(\alpha) + m_i(\alpha)$, so that $\mathcal{H}_i \otimes \mathcal{H}_{E_i^+}^{(\alpha)} \simeq \mathcal{H}_{Q_i^+}^{(\alpha)}$.

At this point, since $U_{Q_i^+}^{(\alpha)}$ acts on the whole of $\mathcal{H}_{Q_i^+}^{(\alpha)}$, how we choose to factor $\mathcal{H}_{Q_i^+}^{(\alpha)}$ to obtain (35) is arbitrary, as long as we choose the factorisation consistently across all α. Recalling that $\mathcal{H}_{Q_i^+}^{(\alpha)} \simeq \mathcal{H}_{Q_i} \otimes \mathcal{H}_{0.i}^{(\alpha)} \simeq \mathcal{H}_i \otimes \mathcal{H}_{E_i^+}^{(\alpha)}$, one possible choice

is to take

$$\dim \mathcal{H}_i^{(\mathrm{out})} = \gcd\Big\{\dim \mathcal{H}_i, \dim \mathcal{H}_{Q_i}\Big\}, \tag{45}$$

$$\dim \mathcal{H}_i^{(\mathrm{in})} = \dim \mathcal{H}_i / \dim \mathcal{H}_i^{(\mathrm{out})}, \tag{46}$$

$$\dim E_i = \dim \mathcal{H}_{Q_i} / \dim \mathcal{H}_i^{(\mathrm{out})}, \tag{47}$$

$$\dim E_{0.i}^{(\alpha)} = \dim \mathcal{H}_{0.i}^{(\alpha)} / \dim H_i^{(\mathrm{in})}. \tag{48}$$

(Note that any of these spaces could turn out to be 1-dimensional.) This choice manifestly satisfies $\dim \mathcal{H}_i = \dim(\mathcal{H}_i^{(\mathrm{out})} \otimes \mathcal{H}_i^{(\mathrm{in})})$, $\dim \mathcal{H}_{Q_i} = \dim(\mathcal{H}_i^{(\mathrm{out})} \otimes E_i)$, and $\dim \mathcal{H}_{0.i}^{(\alpha)} = \dim(\mathcal{H}_i^{(\mathrm{in})} \otimes E_{0.i}^{(\alpha)})$.

To see that this choice is possible for all α, it remains to show that $\dim \mathcal{H}_{0.i}^{(\alpha)}$ is divisible by $\dim \mathcal{H}_i^{(\mathrm{in})}$, so that $\dim E_{0.i}^{(\alpha)}$ is well-defined. First, note that $\dim \mathcal{H}_i^{(\mathrm{in})}$ and $\dim \mathcal{H}_{Q_i}$ are co-prime by (45) and (46). But $\dim \mathcal{H}_i^{(\mathrm{in})}$ divides $\dim \mathcal{H}_{Q_i^+}^{(\alpha)} = \dim \mathcal{H}_{Q_i} \cdot \dim \mathcal{H}_{0.i}^{(\alpha)}$, so $\dim \mathcal{H}_i^{(\mathrm{in})}$ must divide $\dim \mathcal{H}_{0.i}^{(\alpha)}$.

Therefore, we can consistently factor

$$\mathcal{H}_{Q_i} \simeq \mathcal{H}_i^{(\mathrm{out})} \otimes E_i, \tag{49}$$

$$\mathcal{H}_{0.i}^{(\alpha)} \simeq \mathcal{H}_i^{(\mathrm{in})} \otimes E_{0.i}^{(\alpha)}, \tag{50}$$

$$\mathcal{H}_{Q_i^+}^{(\alpha)} \simeq \mathcal{H}_i^{(\mathrm{out})} \otimes \mathcal{H}_i^{(\mathrm{in})} \otimes E_i \otimes E_{0.i}^{(\alpha)}. \tag{51}$$

Recalling that $H_{Q_i^+}^{(\alpha)} \simeq \mathcal{H}_i \otimes \mathcal{H}_{E_i^+}^{(\alpha)}$, we can identify $H_{E_i^+}^{(\alpha)} \simeq E_i \otimes E_{0.i}^{(\alpha)}$, allowing us to rewrite (44) in the form

$$\begin{aligned} U_{Q_0}^\dagger \mathcal{E}(A_i \otimes \mathbb{1}) U_{Q_0} = &\left(\bigoplus_{\alpha : m(\alpha)=0} U_{Q_i^+}^{(\alpha)} \big(A \otimes \mathbb{1}_{E_i} \otimes \mathbb{1}_{E_{0.i}^{(\alpha)}}\big) U_{Q_i^+}^{(\alpha)\,\dagger} \otimes \mathbb{1}_{\mathrm{rest}}^{(\alpha)} \right) \\ &\oplus \left(\bigoplus_{\alpha : n(\alpha)=0} U_{Q_i^+}^{(\alpha)} \big(\bar{A} \otimes \mathbb{1}_{E_i} \otimes \mathbb{1}_{E_{0.i}^{(\alpha)}}\big) U_{Q_i^+}^{(\alpha)\,\dagger} \otimes \mathbb{1}_{\mathrm{rest}}^{(\alpha)} \right) \end{aligned} \tag{52}$$

where A and $\bar{A}$ act on $\mathcal{H}_i^{(\mathrm{in})} \otimes \mathcal{H}_i^{(\mathrm{out})}$.

Finally, note that (52) holds for any single qudit operator $A_i \otimes \mathbb{1}$ on any qudit i. For an arbitrary operator M on n qudits, Theorem 8 follows by expressing M as a real-linear combination of products of single qudit terms, and using additivity, real-linearity and multiplicativity of encodings from Theorem 3(ii). □

An alternative statement of the characterisation in Theorem 8 is given by the following corollary:

Corollary 10 *Let $\mathcal{E} : \mathcal{B}(\bigotimes_{i=1}^n \mathcal{H}_i) \to \mathcal{B}(\bigotimes_{i=1}^{n'} \mathcal{H}'_i)$ be a local encoding with respect to $\{S'_i\}$. Denote $Q_0 = \bigcup_{i,j} S'_i \cap S'_j$ and $Q_i = S'_i \setminus Q_0$ (see Figure 1).*

Then there exist decompositions $\mathcal{H}_{Q_0} \simeq E_0 \otimes (\bigotimes_i \mathcal{H}_i^{(\mathrm{in})})$ *and* $\mathcal{H}_{Q_i} \simeq E_i \otimes \mathcal{H}_i^{(\mathrm{out})}$, *together with identifications* $\mathcal{H}_i \simeq \mathcal{H}_i^{(\mathrm{in})} \otimes \mathcal{H}_i^{(\mathrm{out})}$, *such that the encoding takes the form:*

$$\mathcal{E}(M) = U_{Q_0} \left(\prod_i U_{(i)} \right) \left(M \otimes \mathbb{1} \otimes P_{E_0} + \bar{M} \otimes \mathbb{1} \otimes P_{E_0}^{\perp} \right) \left(\prod_i U_{(i)} \right)^{\dagger} U_{Q_0}^{\dagger} \tag{53}$$

where each unitary $U_{(i)}$ *acts non-trivially only on* $\mathcal{H}_i \otimes E_0 \otimes E_i$, *and the following commutators vanish for all* i, j*:*

$$[U_{(i)}, U_{(j)}] = 0 \quad \textit{and} \quad [P_{E_0}, U_{(i)}] = 0. \tag{54}$$

Proof This is immediate from Theorem 8 and the following definitions of P_{E_0} and $U_{(i)}$:

$$P_{E_0} = \bigoplus_{\alpha : m(\alpha) = 0} \mathbb{1}_{E_0}^{(\alpha)}, \quad \text{and} \quad U_{(i)} = \bigoplus_{\alpha} \left[U_{Q_i^+}^{(\alpha)} \otimes \mathbb{1}_{\mathrm{rest}}^{(\alpha)} \right], \tag{55}$$

where $m(\alpha)$ is defined as in the proof of Theorem 8. □

Theorem 8 characterises what encodings must look like if they are to map local Hamiltonians to local Hamiltonians, and more generally local observables on the original system to local observables on the simulator. We have seen that, because encodings preserve commutators, observables on different qudits of the original system are necessarily mapped to commuting observables on the simulator system, so remain simultaneously measurable.

However, if the subsets S_i' overlap, these observables on the simulator will in general no longer be on disjoint subsets of qudits; tensor products of operators on the original system are not necessarily mapped to tensor products on the simulator. If we impose the additional requirement that tensor products are mapped to tensor products, which is equivalent to requiring that all the subsets S_i' are disjoint, then there is no Q_0 subsytem and the characterisation from Theorem 8 simplifies substantially:

Corollary 11 (Product-preserving encodings)
Let $\mathcal{E} : \mathcal{B}(\bigotimes_{i=1}^{n} \mathcal{H}_i) \to \mathcal{B}(\bigotimes_{i=1}^{n'} \mathcal{H}_i')$ *be a local encoding with respect to* $\{S_i'\}$, *where* S_i' *are disjoint subsets. Then the encoding must take one of the following forms, where* $S_i' = \{i\} \cup E_i$*:*

$$\mathcal{E}(M) = \Big(\bigotimes_i U_{i,E_i} \Big) \Big(M_{1,\ldots,n} \otimes \mathbb{1}_{E_1, E_2, \ldots E_n} \Big) \Big(\bigotimes_i U_{i,E_i}^{\dagger} \Big) \tag{56}$$

or

$$\mathcal{E}(M) = \Big(\bigotimes_i U_{i,E_i} \Big) \Big(\bar{M}_{1,\ldots,n} \otimes \mathbb{1}_{E_1, E_2, \ldots E_n} \Big) \Big(\bigotimes_i U_{i,E_i}^{\dagger} \Big). \tag{57}$$

Thus for tensor products to be mapped to tensor products under encoding, the encoding must be rather trivial. Up to local unitaries, it either consists solely of copies of H, *or* solely of copies of $\bar{M}$; it cannot contain both M and $\bar{M}$. This rules out for example the complex-to-real encoding of Lemma 6.

Corollary 11 applies to product-preserving encodings that map to the entire Hilbert space of the simulator system. We will see shortly that things are more interesting if the local encoding maps into a subspace of the simulator's Hilbert space; non-trivial tensor-product-preserving encodings into a subspace *are* possible.

3.4 Encodings in a subspace

It may be the case that an encoding $\mathcal{E}(H)$ acts only within a subspace S of the simulator system $\mathcal{H}'$. That is, we say a map $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ is an *encoding into the subspace* S if $\mathcal{E}(H)$ has support only on S and the map $H \mapsto \mathcal{E}(H)|_S$ is an encoding. Later we may refer to a map of this form simply as a *subspace encoding* or even just an encoding when the subspace is implicit. We call the subspace $S_\mathcal{E}$ onto which $\mathcal{E}$ maps the *encoded subspace*.

All the conclusions of the above section still hold, but now the target space $S_\mathcal{E}$ is embedded in a larger space $\mathcal{H}'$, so the unitary U is replaced with an isometry V. Any subspace encoding may therefore be written in the form

$$\mathcal{E}(M) = V\left(M^{\oplus p} \oplus \bar{M}^{\oplus q}\right) V^\dagger = V\left(M \otimes P + \bar{M} \otimes Q\right) V^\dagger. \tag{58}$$

We remark that P and Q may be chosen to be any orthogonal projectors on the ancilla system E with $\mathrm{rank}(P) = p$ and $\mathrm{rank}(Q) = q$, provided that the isometry V is changed accordingly. Indeed, even the dimension of the ancilla system E may be increased such that P and Q do not sum to the identity, as long as the map $V|_{\mathrm{supp}(P+Q)}$ is an isometry onto the subspace $S_\mathcal{E}$. This will be useful in the simple characterisation of local subspace encodings given in the next section. Note that $\mathcal{E}(\mathbb{1})$ is the projector onto the subspace $S_\mathcal{E}$.

3.5 Local encodings in a subspace

We can now consider encodings into a subspace that are *local*. Since all the encodings we construct later will not only be local, but in fact will also satisfy the stronger condition of mapping tensor products of operators to tensor products on the simulator, we will restrict our attention here to tensor-product-preserving encodings into a subspace. We therefore want to be able to decompose the simulator system $\mathcal{H}'$ into n subsystems $\mathcal{H}' = \bigotimes_{i=1}^n \mathcal{H}'_i$ such that $\mathcal{H}'_i$ corresponds to $\mathcal{H}_i$ operationally. The encoding of a local observable should then be equivalent to a local observable, in terms of its action on the subspace $S_\mathcal{E}$ into which the encoding maps:

Definition 12 *Let $\mathcal{E} : \mathcal{B}\left(\bigotimes_{i=1}^n \mathcal{H}_j\right) \to \mathcal{B}\left(\bigotimes_{j=1}^n \mathcal{H}'_j\right)$ be a subspace encoding. We say that the encoding is* local *if for any $A_j \in \mathrm{Herm}(\mathcal{H}_j)$, there exists $A'_j \in \mathrm{Herm}(\mathcal{H}'_j)$ such that*

$$\mathcal{E}(A_j \otimes \mathbb{1}) = (A'_j \otimes \mathbb{1})\mathcal{E}(\mathbb{1}). \tag{59}$$

Note that for a simulation of n particles with m particles, this does not mean we require $m = n$, but rather that the m particles can be partitioned into n groups, each of which is labelled by $\mathcal{H}'_j$. First we show that local observables on the original system correspond to local observables on the simulator system:

Proposition 13 *Let $\mathcal{E}$ be a local encoding into the subspace $S_\mathcal{E}$. Let ρ' be a state in the encoded subspace such that $\mathcal{E}(\mathbb{1})\rho' = \rho'$. Let A_j be an observable on qudit j of the original system. Then there exists an observable A'_j on $\mathcal{H}'_j$ such that*

$$\mathrm{Tr}[(A_j \otimes \mathbb{1})\rho] = \mathrm{Tr}[(A'_j \otimes \mathbb{1})\rho'] \tag{60}$$

where $\rho = F(\rho') + B(\rho')$, for F and B defined as

$$F(\rho') = \mathrm{Tr}_E[V^\dagger \rho' V(\mathbb{1} \otimes P)] \text{ and } B(\rho') = \mathrm{Tr}_E[V^\dagger \rho' V(\mathbb{1} \otimes Q)] \tag{61}$$

Proof This is an immediate consequence of Definition 12 and (20). □

It turns out that Definition 12 is equivalent to saying that $\mathcal{E}$ is a tensor product of encodings acting on the the encoded space $S_\mathcal{E}$:

Lemma 14 *An encoding $\mathcal{E}$ is local if and only if it can be written as a "tensor product" of encodings $\varphi_j : \mathrm{Herm}(\mathcal{H}_j) \to \mathrm{Herm}(\mathcal{H}'_j)$ in the following way:*

$$\mathcal{E}\left(\bigotimes_{j=1}^{n} A_j\right) = \left[\bigotimes_{j=1}^{n} \varphi_j(A_j)\right] \mathcal{E}(\mathbb{1}) \tag{62}$$

Proof If there exist encodings φ_j such that (62) holds, then $\mathcal{E}$ is local as for any $A_j \in \mathcal{B}(\mathcal{H}_j)$ one can take $A'_j = \varphi_j(A_j) \in \mathcal{B}(\mathcal{H}'_j)$, and

$$\begin{aligned}
\mathcal{E}(A_j \otimes \mathbb{1}) &= \left[\varphi_j(A_j) \otimes \left(\bigotimes_{k\neq j} \varphi_k(\mathbb{1})\right)\right] \mathcal{E}(\mathbb{1}) && (63)\\
&= \left[\varphi_j(A_j)\varphi_j(\mathbb{1}) \otimes \left(\bigotimes_{k\neq j} \varphi_k(\mathbb{1})\right)\right] \mathcal{E}(\mathbb{1}) && (64)\\
&= \left[(\varphi_j(A_j) \otimes \mathbb{1}) \left(\bigotimes_{k=1}^{n} \varphi_k(\mathbb{1})\right)\right] \mathcal{E}(\mathbb{1}) && (65)\\
&= (A'_j \otimes \mathbb{1})\mathcal{E}(\mathbb{1}). && (66)
\end{aligned}$$

For the converse, we will first show that the map $A_j \mapsto A'_j$ can be taken to be a subspace encoding. Since $A'_j \in \mathrm{Herm}(\mathcal{H}'_j)$ is Hermitian, we have

$$(A'_j \otimes \mathbb{1})\mathcal{E}(\mathbb{1}) = \mathcal{E}(A_j \otimes \mathbb{1}) = \mathcal{E}(A_j \otimes \mathbb{1})^\dagger = \mathcal{E}(\mathbb{1})(A'_j \otimes \mathbb{1}), \tag{67}$$

so $A'_j \otimes \mathbb{1}$ commutes with $\mathcal{E}(\mathbb{1})$.

For a given j, consider the subspace T_j of $\mathcal{H}'_j$ which is entirely annihilated by $\mathcal{E}(\mathbb{1})$, defined by $T_j = \{|\psi\rangle \in \mathcal{H}_j \ : \ (|\psi\rangle\langle\psi| \otimes \mathbb{1})\mathcal{E}(\mathbb{1}) = 0\}$. We will choose to take $\varphi_j(A_j) = \Pi_j A'_j \Pi_j$ where Π_j is the projector onto $T_j^\perp$. We will show that φ_j is a subspace encoding, by showing the requirements of Theorem 3(i) hold in the subspace $T_j^\perp$: Hermiticity preservation, spectrum preservation and real-linearity. First note that $\varphi_j(A_j)$ is Hermitian and has support only on $T_j^\perp$.

The projector $(\mathbb{1} - \Pi_j) \otimes \mathbb{1}$ annihilates $\mathcal{E}(\mathbb{1})$ by definition of T_j, so $(\Pi_j \otimes \mathbb{1})\mathcal{E}(\mathbb{1}) = \mathcal{E}(\mathbb{1})$. Therefore

$$[\varphi_j(A_j) \otimes \mathbb{1}]\mathcal{E}(\mathbb{1}) = [\Pi_j A'_j \Pi_j \otimes \mathbb{1}]\mathcal{E}(\mathbb{1}) = \mathcal{E}(A_j \otimes \mathbb{1}), \tag{68}$$

where we have used the fact that $\mathcal{E}(\mathbb{1})$ commutes with $A'_j \otimes \mathbb{1}$. Thus $\varphi_j(A_j)$ can be used as a replacement for A'_j in (59) which has support only on $T_j^\perp$.

We know that $\varphi_j(A_j) \otimes \mathbb{1}$ commutes with $\mathcal{E}(\mathbb{1})$ and is therefore block diagonal with respect to the $\mathcal{E}(\mathbb{1}), \mathbb{1} - \mathcal{E}(\mathbb{1})$ split. Furthermore since $\varphi_j(A_j)$ has no support on T_j, no eigenvalues of $\varphi_j(A_j) \otimes \mathbb{1}$ are completely annihilated when multiplied by $\mathcal{E}(\mathbb{1})$. Therefore

$$\mathrm{spec}(\varphi_j(A_j)|_{T_j^\perp}) = \mathrm{spec}(\mathcal{E}(A_j \otimes \mathbb{1})|_{S_\mathcal{E}}) = \mathrm{spec}(A_j). \tag{69}$$

Next we show that φ_j is real-linear, using the real-linearity of $\mathcal{E}$. For any $\lambda, \mu \in \mathbb{R}$, and $A_j, B_j \in \mathrm{Herm}(\mathcal{H})$,

$$\begin{aligned}
[\varphi_j(\lambda A_j + \mu B_j) \otimes \mathbb{1}]\mathcal{E}(\mathbb{1}) &= \mathcal{E}((\lambda A_j + \mu B_j) \otimes \mathbb{1}) && (70)\\
&= \lambda\mathcal{E}(A_j \otimes \mathbb{1}) + \mu\mathcal{E}(B_j \otimes \mathbb{1}) && (71)\\
&= [(\lambda\varphi_j(A_j) + \mu\varphi_j(B_j)) \otimes \mathbb{1}]\mathcal{E}(\mathbb{1}) && (72)\\
\Leftrightarrow [(\lambda\varphi_j(A_j) + \mu\varphi_j(B_j) - \varphi_j(\lambda A_j + \mu B_j)) \otimes \mathbb{1}]\mathcal{E}(\mathbb{1}) &= 0. && (73)
\end{aligned}$$

For real-linearity of φ_j we need to show that $M = \lambda\varphi_j(A_j) + \mu\varphi_j(B_j) - \varphi_j(\lambda A_j + \mu B_j)$ vanishes. This follows because $M \otimes \mathbb{1}$ commutes with and is annihilated by $\mathcal{E}(\mathbb{1})$, but M has no support on T_j. Therefore φ_j is an encoding into the subspace $T_j^\perp$.

It remains to show that $\mathcal{E}$ can be written in the form of (62). This follows from the fact that $\mathcal{E}$ and φ_j are Jordan homomorphisms, and $(A_j \otimes \mathbb{1})(\mathbb{1} \otimes B_k) = (\mathbb{1} \otimes B_k)(A_j \otimes \mathbb{1})$. So for example for a bipartite system with two subsystems labelled a and b:

$$\begin{aligned}
\mathcal{E}(A_a \otimes B_b) &= \mathcal{E}(A_a \otimes \mathbb{1})\mathcal{E}(\mathbb{1} \otimes B_b) && (74)\\
&= [\varphi_a(A_a) \otimes \mathbb{1}]\,\mathcal{E}(\mathbb{1})\,[\mathbb{1} \otimes \varphi_b(B_b)]\,\mathcal{E}(\mathbb{1}) && (75)\\
&= [\varphi_a(A_a) \otimes \varphi_b(B_b)]\,\mathcal{E}(\mathbb{1}). && (76)
\end{aligned}$$

□

We remark that if $\mathcal{E}$ and φ_j are extended to homomorphisms on all matrices as described in Theorem 8, then (62) holds for all matrices, not just Hermitian ones. This is because the enveloping algebra for the Hermitian matrices includes all matrices, so any matrix can be written as a product of Hermitian matrices.

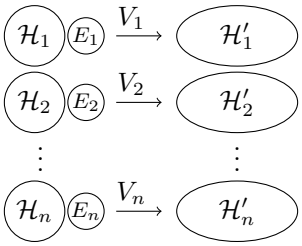


Figure 2: Any local encoding within a subspace can be represented as a tensor product of isometries, as illustrated here.

This extension to all matrices may seem problematic: for example, when calculating $\mathcal{E}(i\mathbb{1})$ one could put the factor of i on any one of the subsystems $\mathcal{H}_j$ before appplying (62). This just implies that the encodings φ_j must satisfy some extra constraints, in order for the overall map to be an encoding.

In fact, we are able to use this condition to derive the following general form of a local encoding (see Figure 2):

Theorem 15 *A map $\mathcal{E} : \mathcal{B}(\bigotimes_{j=1}^n \mathcal{H}_j) \to \mathcal{B}(\bigotimes_{j=1}^n \mathcal{H}'_j)$ is a local encoding if and only if there exist n ancilla systems E_j such that $\mathcal{E}$ is of the form*

$$\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger \tag{77}$$

where

- *V is a local isometry: $V = \bigotimes_j V_j$ for isometries $V_j : \mathcal{H}_j \otimes E_j \to \mathcal{H}'_j$.*
- *P and Q are orthogonal projectors on $E = \bigotimes_j E_j$, and are locally distinguishable: for all j, there exist orthogonal projectors P_{E_j} and Q_{E_j} acting on E_j such that $(P_{E_j} \otimes \mathbb{1})P = P$ and $(Q_{E_j} \otimes \mathbb{1})Q = Q$.*

Proof If $\mathcal{E}$ is of the form given above then by Theorem 3 it is an encoding into the subspace $\mathcal{E}(\mathbb{1}) = V(\mathbb{1} \otimes (P + Q))V^\dagger$. It is easy to check that $\mathcal{E}$ is local: for $A_j \in \mathrm{Herm}(\mathcal{H}_j)$, just take $A'_j = V_j(A_j \otimes P_{E_j} + \bar{A}_j \otimes Q_{E_j})V_j^\dagger \in \mathrm{Herm}(\mathcal{H}'_j)$ and use the conditions of the theorem.

For the converse, note that since $\mathcal{E}$ is an encoding, it must be of the form $\mathcal{E}(M) = W(M \otimes \widetilde{P} + \bar{M} \otimes \widetilde{Q})W^\dagger$, where $\widetilde{P}$ and $\widetilde{Q}$ are projectors on an ancilla system $\widetilde{E}$ and $W : \mathcal{H} \otimes \widetilde{E} \to \mathcal{H}'$ is an isometry. By Lemma 14, there exist n encodings φ_j such that $\mathcal{E}(A_j \otimes \mathbb{1}) = (\varphi_j(A_j) \otimes \mathbb{1})\mathcal{E}(\mathbb{1})$ for any $A_j \in \mathrm{Herm}(\mathcal{H}'_j)$. Since φ_j is an encoding, it must be of the form $\varphi_j(A_j) = V_j(A_j \otimes P_{E_j} + \bar{A}_j \otimes Q_{E_j})V_j^\dagger$ where P_{E_j} and Q_{E_j} are projectors on an ancilla system E_j and $V_j : \mathcal{H}_j \otimes E_j \to \mathcal{H}'_j$ is an isometry.

Let $E = \bigotimes_j E_j$ and define an isometry $V = \bigotimes_j V_j : \mathcal{H} \otimes E \to \mathcal{H}'$. Then by

Lemma 14, for any j and $A_j \in \mathcal{B}(\mathcal{H}_j)$:

$$\mathcal{E}(A_j \otimes \mathbb{1}) = W(A_j \otimes \mathbb{1} \otimes \widetilde{P} + \bar{A}_j \otimes \mathbb{1} \otimes \widetilde{Q})W^\dagger \tag{78}$$

$$= V(A_j \otimes P_{E_j} \otimes \mathbb{1} + \bar{A}_j \otimes Q_{E_j} \otimes \mathbb{1})V^\dagger W(\mathbb{1} \otimes (\widetilde{P} + \widetilde{Q}))W^\dagger \tag{79}$$

Substituting in $A_j = i\mathbb{1}$ in the above expression and matching up the $+i$ and $-i$ eigenspaces implies that

$$V(P_{E_j} \otimes \mathbb{1})V^\dagger W(\mathbb{1} \otimes \widetilde{P})W^\dagger = W(\mathbb{1} \otimes \widetilde{P})W^\dagger \tag{80}$$

$$V(Q_{E_j} \otimes \mathbb{1})V^\dagger W(\mathbb{1} \otimes \widetilde{Q})W^\dagger = W(\mathbb{1} \otimes \widetilde{Q})W^\dagger. \tag{81}$$

We can therefore multiply (78) by $W(\mathbb{1} \otimes \widetilde{P})W^\dagger$ to obtain:

$$W(A_j \otimes \mathbb{1} \otimes \widetilde{P})W^\dagger = V(A_j \otimes \mathbb{1})V^\dagger W(\mathbb{1} \otimes \widetilde{P})W^\dagger \tag{82}$$

implying

$$V^\dagger W(\mathbb{1} \otimes \widetilde{P})(A_j \otimes \mathbb{1}) = (A_j \otimes \mathbb{1})V^\dagger W(\mathbb{1} \otimes \widetilde{P}) \tag{83}$$

Let $\sum_l B_l \otimes C_l$ be the operator Schmidt decomposition of $V^\dagger W(\mathbb{1} \otimes \widetilde{P})$, where $B_l \in \mathcal{B}(\mathcal{H}_j)$ and $C_l : \left(\bigotimes_{k \neq j} \mathcal{H}_k\right) \otimes \widetilde{E} \to \left(\bigotimes_{k \neq j} \mathcal{H}_k\right) \otimes E$. Then from (83) we have

$$\sum_l [B_l, A_j] \otimes C_l = 0 \tag{84}$$

which implies $[B_l, A_j] = 0$ for all l by linear independence of the C_l. This holds for all matrices on $A_j \in \mathcal{B}(\mathcal{H}_j)$. So by Schur's lemma, each B_l, and hence also $V^\dagger W(\mathbb{1} \otimes \widetilde{P})$, must act trivially (i.e. as a multiple of the identity) on $\mathcal{H}_j$ for all j, and hence on $\mathcal{H}$.

By the same argument $V^\dagger W(\mathbb{1} \otimes \widetilde{Q})$ acts trivially on all of $\mathcal{H}$ and so we can conclude there must exist an isometry $U : \widetilde{E} \to E$ such that

$$V^\dagger W(\mathbb{1} \otimes \widetilde{P}) = (\mathbb{1} \otimes U\widetilde{P}) \text{ and } V^\dagger W(\mathbb{1} \otimes \widetilde{Q}) = (\mathbb{1} \otimes U\widetilde{Q}) \tag{85}$$

Define $P = U\widetilde{P}U^\dagger$ and $Q = U\widetilde{Q}U^\dagger$, and remember that $\mathcal{E}(M)$ must be in the range of the isometry V by Lemma 14, so we have

$$\mathcal{E}(M) = VV^\dagger \mathcal{E}(M) VV^\dagger = VV^\dagger W(M \otimes \widetilde{P} + \bar{M} \otimes \widetilde{Q})W^\dagger VV^\dagger \tag{86}$$

$$= V(M \otimes P + \bar{M} \otimes Q)V^\dagger \tag{87}$$

and note that (80) implies that $(P_{E_j} \otimes \mathbb{1})P = P$ and $(Q_{E_j} \otimes \mathbb{1})Q = Q$ as required. □

When $\mathcal{E}$ is a local encoding from n qudits to m qudits of the same local dimension d, the space $\mathcal{H}'_j$ is a group of k_j qudits. As described at the end of Section 3.4, the dimension of the ancilla E_j can be increased until it is of size $d^{k_j - 1}$ so that the dimensions of $\mathcal{H}_j \otimes E_j$ and $\mathcal{H}'$ match. If this is done for all j, then all the V_j (and hence also $V = \bigotimes V_j$) are unitaries.

3.6 Composition and approximation of encodings

In this section, we collect some straightforward technical lemmas about encodings which we will need later: that encodings compose properly, and that approximations to encodings behave as one would expect.

Lemma 16 *If $\mathcal{E}_1$ and $\mathcal{E}_2$ are encodings, then their composition $\mathcal{E}_1 \circ \mathcal{E}_2$ is also an encoding. Furthermore, if $\mathcal{E}_1$ and $\mathcal{E}_2$ are both local, then their composition $\mathcal{E}_1 \circ \mathcal{E}_2$ is local.*

Proof By the definition of encodings, we can write

$$\mathcal{E}_1(M) = V(M \otimes P^{(1)} + \bar{M} \otimes Q^{(1)})V^\dagger \tag{88}$$

$$\mathcal{E}_2(M) = W(M \otimes P^{(2)} + \bar{M} \otimes Q^{(2)})W^\dagger \tag{89}$$

for isometries V and W, and orthogonal pairs of projectors $P^{(1)}, Q^{(1)}$ and $P^{(2)}, Q^{(2)}$. Then

$$(\mathcal{E}_1 \circ \mathcal{E}_2)(M) = V\Big[W(M \otimes P^{(2)} + \bar{M} \otimes Q^{(2)})W^\dagger \otimes P^{(1)} \tag{90}$$

$$+\overline{W(M \otimes P^{(2)} + \bar{M} \otimes Q^{(2)})W^\dagger} \otimes Q^{(1)}\Big] V^\dagger \tag{91}$$

$$= U\Big[M \otimes \Big(P^{(2)} \otimes P^{(1)} + \bar{Q}^{(2)} \otimes Q^{(1)}\Big) \tag{92}$$

$$+\bar{M} \otimes \Big(Q^{(2)} \otimes P^{(1)} + \bar{P}^{(2)} \otimes Q^{(1)}\Big)\Big] U^\dagger \tag{93}$$

where $U = V\left(W \otimes P^{(1)} + \bar{W} \otimes Q^{(1)} + \mathbb{1} \otimes (\mathbb{1} - P^{(1)} - Q^{(1)})\right) V^\dagger$ is an isometry. Then observing that $P = P^{(2)} \otimes P^{(1)} + \bar{Q}^{(2)} \otimes Q^{(1)}$ and $Q = Q^{(2)} \otimes P^{(1)} + \bar{P}^{(2)} \otimes Q^{(1)}$ are orthogonal projectors, we conclude that $\mathcal{E}_1 \circ \mathcal{E}_2$ is an encoding.

If $\mathcal{E}_1$ and $\mathcal{E}_2$ are both local then the projectors are locally distinguishable, which means there exist projectors $P^{(a)}_{E_i^{(a)}}$ and $Q^{(a)}_{E_i^{(a)}}$ for $a \in \{1, 2\}$ such that

$$\left(P^{(a)}_{E_i^{(a)}} \otimes \mathbb{1}\right) P^{(a)} = P^{(a)} \quad \text{and} \quad \left(Q^{(a)}_{E_i^{(a)}} \otimes \mathbb{1}\right) Q^{(a)} = Q^{(a)}. \tag{94}$$

We can show that P and Q are locally distinguishable by defining orthogonal projectors on the systems $E_i = E_i^{(2)} \otimes E_i^{(1)}$ as follows:

$$P_{E_i} = P^{(2)}_{E_i^{(2)}} \otimes P^{(1)}_{E_i^{(1)}} + \bar{Q}^{(2)}_{E_i^{(1)}} \otimes Q^{(1)}_{E_i^{(1)}} \text{ and } Q_{E_i} = Q^{(2)}_{E_i^{(2)}} \otimes P^{(1)}_{E_i^{(1)}} + \bar{P}^{(2)}_{E_i^{(1)}} \otimes Q^{(1)}_{E_i^{(1)}} \tag{95}$$

such that $(P_{E_i} \otimes \mathbb{1})P = P$ and $(Q_{E_i} \otimes \mathbb{1})Q = Q$.

Furthermore, since $\mathcal{E}_1$ and $\mathcal{E}_2$ are local, the isometries V and W are tensor products $V = \bigotimes_i V_i$ and $W = \bigotimes_i W_i$, and we can define a local isometry

$$U' = \bigotimes_i V_i \left(W_i \otimes P^{(1)}_{E_i^{(1)}} + \bar{W}_i \otimes Q^{(1)}_{E_i^{(1)}} + \mathbb{1} \otimes (\mathbb{1} - P^{(1)}_{E_i^{(1)}} - Q^{(1)}_{E_i^{(1)}})\right) V_i^\dagger \tag{96}$$

such that $(\mathcal{E}_1 \circ \mathcal{E}_2)(M) = U'(M \otimes P + \bar{M} \otimes Q)U'^\dagger$. □

Next we show that, unsurprisingly, if two encodings are close, the results of applying the encodings to the same operator are also close; and similarly that if two operators are close, the results of applying the same encoding to the operators are close. We first prove a small technical lemma, which will be useful both here and throughout the paper.

Lemma 17 *Let $A, B : \mathcal{H} \to \mathcal{H}'$ and $C : \mathcal{H} \to \mathcal{H}$ be linear maps. Let $\|\cdot\|_a$ be the trace norm or operator norm. Then*

$$\|ACA^\dagger - BCB^\dagger\|_a \leq (\|A\| + \|B\|)\|A - B\|\|C\|_a. \tag{97}$$

Proof The proof is a simple application of the triangle inequality followed by submultiplicativity:

$$\begin{aligned}\|ACA^\dagger - BCB^\dagger\|_a &\leq \|ACA^\dagger - BCA^\dagger\|_a + \|BCA^\dagger - BCB^\dagger\|_a &(98)\\ &\leq \|A - B\|\|C\|_a\|A^\dagger\| + \|B\|\|C\|_a\|A^\dagger - B^\dagger\| &(99)\\ &= (\|A\| + \|B\|)\|A - B\|\|C\|_a &(100)\end{aligned}$$

where we have also used $\|A\| = \|A^\dagger\|$. □

Lemma 18 *Consider two encodings $\mathcal{E}$ and $\widetilde{\mathcal{E}}$ defined by $\mathcal{E}(M) = V(M^{\oplus p} \oplus \bar{M}^{\oplus q})V^\dagger$, $\widetilde{\mathcal{E}}(M) = \widetilde{V}(M^{\oplus p} \oplus \bar{M}^{\oplus q})\widetilde{V}^\dagger$, for some isometries V, $\widetilde{V}$. Then, for any operators M and $\widetilde{M}$:*

(i). $\|\mathcal{E}(M) - \widetilde{\mathcal{E}}(M)\| \leq 2\|V - \widetilde{V}\|\|M\|$;

(ii). $\|\mathcal{E}_{\text{state}}(M) - \widetilde{\mathcal{E}}_{\text{state}}(M)\|_1 \leq 2\|V - \widetilde{V}\|\|M\|_1$;

(iii). $\|\mathcal{E}(M) - \mathcal{E}(\widetilde{M})\| = \|M - \widetilde{M}\|$.

Proof Write $M' = M^{\oplus p} \oplus \bar{M}^{\oplus q}$. Then, for the first part,

$$\|\mathcal{E}(M) - \widetilde{\mathcal{E}}(M)\| = \|VM'V^\dagger - \widetilde{V}M'\widetilde{V}^\dagger\| \leq 2\|V - \widetilde{V}\|\|M\| \tag{101}$$

by Lemma 17, using $\|M'\| = \|M\|$. For the second part, recall that $\mathcal{E}_{\text{state}}(\rho)$ is either defined as $V(\rho \otimes \sigma)V^\dagger$ or $V(\bar{\rho} \otimes \sigma)V^\dagger$, dependent on whether $p \geq 1$, for some fixed state σ. Then, writing $M' = M \otimes \sigma$ or $M' = \bar{M} \otimes \sigma$ and observing that $\|M'\|_1 = \|M\|_1$, the argument is the same as the first part (replacing the operator norm with the trace norm appropriately).

The third part is essentially immediate:

$$\|\mathcal{E}(M) - \mathcal{E}(\widetilde{M})\| = \|V((M - \widetilde{M})^{\oplus p} \oplus (\bar{M} - \bar{\widetilde{M}})^{\oplus q})V^\dagger\| = \|M - \widetilde{M}\|. \tag{102}$$

□

4 Hamiltonian simulation

4.1 Perfect simulation

We have seen that encodings capture the notion of one Hamiltonian exactly reproducing all the physics of another. We will be interested in a less restrictive notion, where this holds only for the low-energy part of the first Hamiltonian. This concept can be captured by generalising the idea of encodings to *simulations*. Let $H \in \mathcal{B}((\mathbb{C}^d)^{\otimes n})$ and $H' \in \mathcal{B}((\mathbb{C}^{d'})^{\otimes m})$ for some $m \geq n$. We usually think of the local dimensions d, d' as fixed, but the number of qudits n, m as growing. Recall that $S_{\leq\Delta(H')} = \text{span}\{|\psi\rangle : H|\psi\rangle = \lambda|\psi\rangle, \lambda \leq \Delta\}$ denotes the low energy space of H' and $P_{\leq\Delta(H')}$ denotes the projector onto this space.

Definition 19 *We say that H' perfectly simulates H below energy Δ if there is a local encoding $\mathcal{E}$ into the subspace $S_\mathcal{E}$ such that:*

(i). $S_\mathcal{E} = S_{\leq\Delta(H')}$ (or equivalently $\mathcal{E}(\mathbb{1}) = P_{\leq\Delta(H')}$);

(ii). $H'|_{\leq\Delta} = \mathcal{E}(H)|_{S_\mathcal{E}}$.

Note that condition (i) is crucial in order for it to make sense to compare $H'|_{\leq\Delta}$ and $\mathcal{E}(H)|_{S_\mathcal{E}}$. When condition (i) holds, condition (ii) is equivalent to $H'_{\leq\Delta} = \mathcal{E}(H)$, where $H'_{\leq\Delta} = H'P_{\leq\Delta(H')}$ is the low energy part of H'.

To gain some intuition for the above definition, taking $\mathcal{E}$ to be the identity map, we see that H perfectly simulates itself. Further, for any $U \in U(d)$, we see that $U^{\otimes n}H(U^\dagger)^{\otimes n}$ is a perfect simulation of H. This freedom to apply local unitaries allows us, for example, to relabel Pauli matrices in the Pauli expansion of H. It also allows us to bring 2-qubit interactions into a canonical form [CM16]. Imagine we have a Hamiltonian on n qubits which can be written as a sum of 2-local terms, each proportional to some 2-qubit interaction H which is symmetric under interchange of the qubits. Then it is not hard to show [CM16] that there exists $U \in SU(2)$ such that

$$U^{\otimes 2}H(U^\dagger)^{\otimes 2} = \sum_{s\in\{x,y,z\}} \alpha_s \sigma_s \otimes \sigma_s + \sum_{t\in\{x,y,z\}} \beta_t(\sigma_t \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_t) \qquad (103)$$

for some weights $\alpha_s, \beta_t \in \mathbb{R}$. Applying $U^{\otimes n}$ to the whole Hamiltonian simulates the H interactions with interactions of this potentially simpler form.

Both of these examples of perfect simulations are actually also encodings. As an example of a perfect simulation which is not an encoding, we observe that qubit Hamiltonians can simulate qudit Hamiltonians.

Lemma 20 *Let H be a k-local qudit Hamiltonian on n qudits with local dimension d. Then, for any $\Delta \geq \|H\|$, there is a $k\lceil\log_2 d\rceil$-local qubit Hamiltonian H' which perfectly simulates H below energy Δ.*

Proof We use a local encoding $\mathcal{E}(M) = VMV^\dagger$, where $V = W^{\otimes n}$, and $W : \mathbb{C}^d \to (\mathbb{C}^2)^{\otimes\lceil\log_2 d\rceil}$ is an arbitrary isometry. Write $P = \mathbb{1} - WW^\dagger$ for the

projector onto the subspace orthogonal to the image of W (if d is a power of 2, $P = 0$). Then we define the Hamiltonian

$$H' = \mathcal{E}(H) + \Delta' \sum_{i=1}^{n} P_i, \tag{104}$$

for some $\Delta' > \Delta$. The nullspace of the positive semidefinite operator $P := \sum_{i=1}^{n} P_i$ is precisely the image of V, and the smallest nonzero eigenvalue of P is Δ'. So, as $\Delta < \Delta'$, $\mathcal{E}$ is an encoding into the subspace $S_{\mathcal{E}} = S_{\leq \Delta(H')}$; and as $\Delta \geq \|H\|$, $H'|_{\leq \Delta} = \mathcal{E}(H)|_{S_{\mathcal{E}}}$. Thus H' meets the requirements of Definition 19 and perfectly simulates H below energy Δ. □

Another case where we can achieve perfect simulation is the simulation of complex Hamiltonians with real Hamiltonians, using an alternative to the complex-to-real encoding of Lemma 6 where no single qubit corresponds to the ancilla qubit of Lemma 6. This enables us to make the subspace encoding in the simulation local.

Lemma 21 *For any integer k, let H be a k-local qubit Hamiltonian. Then for any $\Delta \geq 2\|H\|$ there is a real $2k$-local qubit Hamiltonian H' which simulates H perfectly below energy Δ.*

Proof Let H be a k-local qubit Hamiltonian, and let $h = \bigotimes_{i=1}^{k} \sigma_{s_i}$ with $s_i \in \{x, y, z\}$ be a k-local term in the Pauli decomposition of H. The complex-to-real encoding φ from Lemma 6 maps individual Paulis as follows:

$$\varphi(\mathbb{1}) = \mathbb{1} \oplus \mathbb{1} \tag{105}$$

$$\varphi(\sigma_{x,z}) = \sigma_{x,z} \oplus \sigma_{x,z} = \mathbb{1} \otimes \sigma_{x,z} \tag{106}$$

$$\varphi(\sigma_y) = J(\sigma_y \oplus \sigma_y) = \sigma_y \otimes \sigma_y. \tag{107}$$

For each qubit j in the original Hamiltonian H, add an additional qubit labelled j' and apply the map φ separately to these pairs of qubits. This results in a term h' on $2k$ qubits of the following form:

$$h' = \bigotimes_{j=1}^{k} \left(|+_y\rangle\langle+_y|_{j'} \otimes \sigma_{s_j} + |-_y\rangle\langle-_y|_{j'} \otimes \bar{\sigma}_{s_j} \right) \tag{108}$$

Restricted to the space S spanned by $|+_y\rangle^{\otimes n}$ and $|-_y\rangle^{\otimes n}$ on the ancilla qubits, h is of the desired form. (Indeed, the restriction recovers the complex-to-real encoding of Lemma 6.) Let $\widetilde{H}$ be the total Hamiltonian formed by the sum of the h' terms. Then $\widetilde{H}$ is real and

$$\widetilde{H}|_S = |+_y\rangle\langle+_y|^{\otimes n} \otimes H + |-_y\rangle\langle-_y|^{\otimes n} \otimes \bar{H} \tag{109}$$

We can add a term $\Delta' H_0$ where $\Delta' > \Delta$ and $H_0 = \sum_i (Y_{i'} Y_{(i+1)'} + \mathbb{1})$ is zero on S and is $\geq \mathbb{1}$ on $S^{\perp}$. The overall Hamiltonian $H' = \widetilde{H} + \Delta' H_0$ is therefore real and, since $\Delta' > \Delta \geq 2\|H\|$, $S_{\leq \Delta(H')} = S$ and $H'|_{\leq \Delta} = \widetilde{H}|_S$. □

4.2 Approximate simulation

In general we may not be able to achieve perfect simulation, so it is natural to generalise this concept to allow approximate simulations. If condition (i) in Definition 19 no longer holds exactly for a map $\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$, it is not immediately clear how to generalise condition (ii), as $H'_{\leq\Delta}$ and $\mathcal{E}(H)$ now have support on different spaces. However, if condition (i) holds approximately such that $\|\mathcal{E}(\mathbb{1}) - P_{\leq\Delta(H')}\| \leq \eta$, then there exists an alternative encoding $\widetilde{\mathcal{E}}(M) = \widetilde{V}(M \otimes P + \bar{M} \otimes Q)\widetilde{V}^\dagger$ such that $\|\widetilde{V} - V\| \leq \sqrt{2}\eta$ and $\widetilde{\mathcal{E}}(\mathbb{1}) = P_{\leq\Delta(H')}$ (see Lemma 23 below); so we can compare $H'_{\leq\Delta}$ and $\widetilde{\mathcal{E}}(H)$.

Definition 22 *We say that H' is a (Δ, η, ϵ)-simulation of H if there exists a local encoding $\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$ such that:*

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

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

We say that a family $\mathcal{F}'$ of Hamiltonians can simulate a family $\mathcal{F}$ of Hamiltonians if, for any $H \in \mathcal{F}$ and any $\eta, \epsilon > 0$ and $\Delta \geq \Delta_0$ (for some $\Delta_0 > 0$), there exists $H' \in \mathcal{F}'$ such that H' is a (Δ, η, ϵ)-simulation of H. We say that the simulation is efficient if, in addition, for H acting on n qudits, $\|H'\| = \mathrm{poly}(n, 1/\eta, 1/\epsilon, \Delta)$; H' is efficiently computable given H, Δ, η and ϵ; and each local isometry V_i in the decomposition of V from Theorem 15 is itself a tensor product of isometries which map to $O(1)$ qudits.

We usually think of Δ as satisfying $\Delta \gg \|H\|$. But we can also consider smaller Δ by only simulating H up to some energy cutoff. We may interpret Definition 22 as stating that $H'_{\leq\Delta}$ is close to an encoding $\widetilde{\mathcal{E}}(H)$ of H, and that the encoding map $\widetilde{\mathcal{E}}$ is close to a local encoding $\mathcal{E}$. However, we assume that $\mathcal{E}$ is the map that we understand and have access to, whereas all we know about $\widetilde{\mathcal{E}}$ is that it exists.

A perfect simulation of H by H' below energy Δ is a (Δ, η, ϵ)-simulation of H with $\eta = \epsilon = 0$. Observe that every local encoding is a perfect simulation with $\Delta = \infty$. Reducing the inaccuracy η, ϵ of the simulation will typically require expending more "effort", e.g. by increasing the strength of the local interactions.

An alternative definition might try to compare $H'_{\leq\Delta}$ and $\mathcal{E}(H)$ even though they have different support. This would be essentially equivalent to our definition because, from Lemma 18 and the reverse triangle inequality,

$$\left| \|H'_{\leq\Delta} - \mathcal{E}(H)\| - \|H'_{\leq\Delta} - \widetilde{\mathcal{E}}(H)\| \right| \leq 2\|V - \widetilde{V}\| \|H^{\oplus p} \oplus \bar{H}^{\oplus q}\| \leq 2\eta\|H\|. \quad (110)$$

Thus the two definitions are equivalent up to a $O(\eta\|H\|)$ term. Our simulations will in general assume that $\eta = O(1/\mathrm{poly}(\|H\|))$, making the difference negligible. It is also worth noting that this alternative definition appears to result in worse bounds in Lemma 24 and Corollary 29 below.

We remark that our physically motivated definition of simulation is very similar to one previously introduced by Bravyi and Hastings [BH17]. The main differences are:

(i). The second part of the definition in [BH17] is stated as

$$\|H - \widetilde{V}^\dagger H' \widetilde{V}\| \leq \epsilon. \tag{111}$$

But we have $\|H - \widetilde{V}^\dagger H' \widetilde{V}\| = \|\widetilde{V} H \widetilde{V}^\dagger - \widetilde{V}\widetilde{V}^\dagger H' \widetilde{V}\widetilde{V}^\dagger\| = \|\widetilde{V} H \widetilde{V}^\dagger - H'_{\leq \Delta}\|$, which matches the term $\|H'_{\leq\Delta} - \widetilde{\mathcal{E}}(H)\|$ in our definition, except that our encoding $\widetilde{\mathcal{E}}(H)$ may be of the more general form $\widetilde{V}(H^{\oplus p} \oplus \bar{H}^{\oplus q})\widetilde{V}^\dagger$. As discussed above, this is essential to enable e.g. complex Hamiltonians to be encoded as real Hamiltonians.

(ii). We insist that $\mathcal{E}$ is local, whereas [BH17] deliberately does not impose any restriction on the isometry V, other than to say it should be sufficiently simple in practice. This enables us to find stronger implications of our notion of simulation for error-tolerance and computational complexity.

We also remark that, although Definition 22 requires simulation in the low-energy subspace, this can readily be generalised to other types of subspace, by replacing $P_{\leq\Delta(H')}$ by a projector onto the subspace of interest. However, some of the physical consequences of Definition 22 later in this section do depend on the simulation being in the low-energy subspace. All the simulations we construct will achieve this.

We now prove the previously promised claim that if the isometry V used in an encoding approximately maps to the ground space of H', there exists an isometry $\widetilde{V}$ close to V which maps exactly to this ground space. See [BH17] for a similar result.

Lemma 23 *Let $\mathcal{E} : \mathcal{B}(\mathcal{H}) \to \mathcal{B}(\mathcal{H}')$ be a subspace encoding of the form $\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$, and let Π be the projector onto a subspace $S \subseteq \mathcal{H}'$. If $\|\Pi - \mathcal{E}(\mathbb{1})\| < 1$, then there exists an isometry $\widetilde{V} : \mathcal{H} \to \mathcal{H}'$ such that $\|\widetilde{V} - V\| \leq \sqrt{2}\|\Pi - \mathcal{E}(\mathbb{1})\|$ and the corresponding encoding $\widetilde{\mathcal{E}}(M) = \widetilde{V}(M \otimes P + \bar{M} \otimes Q)\widetilde{V}^\dagger$ satisfies $\widetilde{\mathcal{E}}(\mathbb{1}) = \Pi$.*

Proof Recall that $\mathcal{E}(\mathbb{1})$ is a projector. If $\|\Pi - \mathcal{E}(\mathbb{1})\| < 1$, then $\mathrm{rank}(\Pi) = \mathrm{rank}(\mathcal{E}(\mathbb{1}))$ and hence there exists a unitary U on $\mathcal{H}'$ such that $\Pi = U\mathcal{E}(\mathbb{1})U^\dagger$. One can show using Jordan's lemma that U can be chosen to obey the bound $\|U - \mathbb{1}\| \leq \sqrt{2}\|\Pi - \mathcal{E}(\mathbb{1})\|$; the short argument is contained in the proof of Lemma 3 in [BH17].

Defining $\widetilde{V} = UV$, we have $\widetilde{\mathcal{E}}(\mathbb{1}) = U\mathcal{E}(\mathbb{1})U^\dagger = \Pi$ and

$$\|\widetilde{V} - V\| \leq \|U - \mathbb{1}\|\|V\| \leq \sqrt{2}\|\Pi - \mathcal{E}(\mathbb{1})\| \tag{112}$$

as desired. □

Importantly, the notion of simulation we use is transitive: if A simulates B, and B simulates C, then A simulates C. We now formalise this as a lemma; a very similar result to this was shown by Bravyi and Hastings [BH17], but as our encodings are somewhat more general to those they consider we include a proof.

Lemma 24 *Let A, B, C be Hamiltonians such that A is a $(\Delta_A, \eta_A, \epsilon_A)$-simulation of B and B is a $(\Delta_B, \eta_B, \epsilon_B)$-simulation of C. Suppose $\epsilon_A, \epsilon_B \leq \|C\|$ and $\Delta_B \geq \|C\| + 2\epsilon_A + \epsilon_B$. Then A is a (Δ, η, ϵ)-simulation of C, where $\Delta \geq \Delta_B - \epsilon_A$,*

$$\eta = \eta_A + \eta_B + O\left(\frac{\epsilon_A}{\Delta_B - \|C\| + \epsilon_B}\right) \quad \text{and} \quad \epsilon = \epsilon_A + \epsilon_B + O\left(\frac{\epsilon_A \|C\|}{\Delta_B - \|C\| + \epsilon_B}\right). \tag{113}$$

Note that any good simulation should satisfy $\Delta_B \gg \|C\|$ (see Proposition 27 below for one reason why) in which case the condition on Δ_B is easily satisfied and we have $\eta = \eta_A + \eta_B + o(1)$ and $\epsilon \approx \epsilon_A + \epsilon_B$.

Proof We closely follow the argument of [BH17, Lemma 3]. Let $\mathcal{E}_A$ be the local encoding corresponding to the simulation of B with A, and let $\mathcal{E}_B$ be the local encoding corresponding to the simulation of C with B. We will use the composed map $\mathcal{E} = \mathcal{E}_A \circ \mathcal{E}_B$ to simulate C with A. By Lemma 16, this map is indeed a local encoding.

Let V_A and V_B be the isometries in the definition of $\mathcal{E}_A$ and $\mathcal{E}_B$. Recall from the definition of simulation that there exist isometries $\widetilde{V}_A$, $\widetilde{V}_B$ such that $\|\widetilde{V}_A - V_A\| \leq \eta_A$, $\|\widetilde{V}_B - V_B\| \leq \eta_B$, $\widetilde{V}_A \widetilde{V}_A^\dagger = P_{\leq \Delta_A(A)}$, $\widetilde{V}_B \widetilde{V}_B^\dagger = P_{\leq \Delta_B(B)}$. We define the encodings $\widetilde{\mathcal{E}}_A$, $\widetilde{\mathcal{E}}_B$ to be the encodings obtained by replacing V_A with $\widetilde{V}_A$ and V_B with $\widetilde{V}_B$. Note that composing these maps to obtain $\widetilde{\mathcal{E}}_A \circ \widetilde{\mathcal{E}}_B$ makes sense ($\widetilde{\mathcal{E}}_B$ maps C to the low-energy part of B, and $\widetilde{\mathcal{E}}_A$ maps all of B to the low-energy part of A).

Let N be the dimension of $S_{\leq \Delta_B(B)}$. By Lemma 26, the Nth smallest eigenvalue of B is bounded by $\lambda_N(B) \leq \|C\| + \epsilon_B$. Therefore the condition $\Delta_B \geq \|C\| + 2\epsilon_A + \epsilon_B$ allows us to put a lower bound on Δ_G, the spectral gap between the Nth and $(N+1)$th eigenvalues of B:

$$\Delta_G = \lambda_{N+1}(B) - \lambda_N(B) > \Delta_B - \|C\| - \epsilon_B \geq 2\epsilon_A. \tag{114}$$

Let $\widetilde{\mathcal{E}}_A(B) = \widetilde{V}_A(B^{\oplus p} \oplus \bar{B}^{\oplus q})\widetilde{V}_A^\dagger$. By Lemma 26, $\lambda_{N(p+q)}(A) \leq \lambda_N(B) + \epsilon_A$ and $\lambda_{N(p+q)+1}(A) \geq \lambda_{N+1}(B) - \epsilon_A$, so the condition $\Delta_G > 2\epsilon_A$ implies that there exists Δ such that $\lambda_{N(p+q)}(A) < \Delta < \lambda_{N(p+q)+1}(A)$. Furthermore, since $\lambda_{N(p+q)+1}(A) \geq \lambda_{N+1}(B) - \epsilon_A > \Delta_B - \epsilon_A$, we can choose Δ to be at least as big as $\Delta_B - \epsilon_A$.

Let $B' = B^{\oplus p} \oplus \bar{B}^{\oplus q}$, so we can write $\widetilde{\mathcal{E}}_A(B) = \widetilde{V}_A B' \widetilde{V}_A^\dagger$. It is shown in the proof of [BH17, Lemma 3] that there exists a unitary operator U such that

$$S_{\leq \Delta(A)} = U \widetilde{V}_A S_{\leq \Delta_B(B')} \tag{115}$$

and $\|U - \mathbb{1}\| \leq 2\sqrt{2}\epsilon_A/\Delta_G$. That is, $U\widetilde{V}_A$ maps the low-energy subspace of B' precisely onto the low-energy subspace of A. Note that the existence of such a

U is nontrivial, as all we know in advance from the fact that A simulates B is that $\widetilde{V}_A$ maps all of B' into the less low-energy subspace $S_{\leq \Delta_A(A)}$.

The composed approximate encoding in the simulation of C by A will be $\widetilde{\mathcal{E}}(M) = U\widetilde{\mathcal{E}}_A(\widetilde{\mathcal{E}}_B(M))U^\dagger$. By (115), $\widetilde{\mathcal{E}}$ maps the Hilbert space of C onto $S_{\leq \Delta(A)}$. The overall isometry $\widetilde{V}$ in the encoding $\widetilde{\mathcal{E}}_A \circ \widetilde{\mathcal{E}}_B$ is obtained from the isometry V in the encoding $\mathcal{E}$ by replacing V_A with $\widetilde{V}_A$ and V_B with $\widetilde{V}_B$. By the triangle inequality and Lemma 16, $\|V - \widetilde{V}\| \leq \eta_A + \eta_B$, so

$$\eta = \|V - U\widetilde{V}\| \leq \eta_A + \eta_B + O(\epsilon_A \Delta_G^{-1}). \tag{116}$$

Therefore, $\mathcal{E}$ meets condition (i) from Definition 22 for simulation of C with A.

It remains to show condition (ii). We aim to bound $\|A_{\leq\Delta} - U\widetilde{\mathcal{E}}_A(\widetilde{\mathcal{E}}_B(C))U^\dagger\|$, which, by the triangle inequality, is upper-bounded by

$$\begin{aligned}&\|A_{\leq\Delta} - U\widetilde{\mathcal{E}}_A(\widetilde{\mathcal{E}}_B(C))U^\dagger\| \\ &\quad \leq \|A_{\leq\Delta} - U\widetilde{\mathcal{E}}_A(B_{\leq\Delta_B})U^\dagger\| + \|U\widetilde{\mathcal{E}}_A(B_{\leq\Delta_B})U^\dagger - U\widetilde{\mathcal{E}}_A(\widetilde{\mathcal{E}}_B(C))U^\dagger\|.\end{aligned} \tag{117}$$

The second term in (117) is precisely equal to $\|B_{\leq\Delta_B} - \widetilde{\mathcal{E}}_B(C)\|$. By the assumption of the present lemma that B is a $(\Delta_B, \eta_B, \epsilon_B)$-simulation of C, this term is upper-bounded by ϵ_B. In order to deal with the first term in (117), we rewrite it as

$$\|A_{\leq\Delta}U\widetilde{V}_A - U\widetilde{V}_A B'_{\leq\Delta_B}\|. \tag{118}$$

We write $U = \mathbb{1} + M$, so

$$A_{\leq\Delta}U\widetilde{V}_A - U\widetilde{V}_A B'_{\leq\Delta_B} = P_{\leq\Delta(A)}(A_{\leq\Delta}U\widetilde{V}_A - U\widetilde{V}_A B'_{\leq\Delta_B})P_{\leq\Delta_B(B')} \tag{119}$$

$$= P_{\leq\Delta(A)}(A\widetilde{V}_A - \widetilde{V}_A B')P_{\leq\Delta_B(B')} \tag{120}$$

$$+ A_{\leq\Delta}M\widetilde{V}_A P_{\leq\Delta_B(B')} - P_{\leq\Delta(A)}M\widetilde{V}_A B'_{\leq\Delta_B}. \tag{121}$$

For the first part,

$$\|P_{\leq\Delta(A)}(A\widetilde{V}_A - \widetilde{V}_A B')P_{\leq\Delta_B(B')}\| \leq \|A\widetilde{V}_A - \widetilde{V}_A B'\| = \|A_{\leq\Delta_A} - \widetilde{V}_A B'\widetilde{V}_A^\dagger\| \leq \epsilon_A \tag{122}$$

by simulation of B with A. The second part is bounded by $\|M\|\|A_{\leq\Delta}\|$ and the third by $\|M\|\|B'_{\leq\Delta_B}\|$. We have $\|M\| = O(\epsilon_A\Delta_G^{-1})$ by (115). By simulation of B with A and (115), $\|A_{\leq\Delta}\| \leq \|B'_{\leq\Delta_B}\| + \epsilon_A$; by simulation of C with B, $\|B'_{\leq\Delta_B}\| = \|B_{\leq\Delta_B}\| \leq \|C\| + \epsilon_B$. Combining all the terms, we get the overall bound that

$$\|A_{\leq\Delta} - U\widetilde{\mathcal{E}}_A(\widetilde{\mathcal{E}}_B(C))U^\dagger\| \leq \epsilon_A + \epsilon_B + 2\sqrt{2}\epsilon_A\Delta_G^{-1}(\|C\| + \epsilon_A + 2\epsilon_B). \tag{123}$$

Since $\epsilon_A, \epsilon_B \leq \|C\|$ and $\Delta_B \leq \Delta_G + \|C\| + \epsilon_B$, we have that the overall error ϵ is

$$\epsilon = \epsilon_A + \epsilon_B + O\left(\frac{\epsilon_A\|C\|}{\Delta_B - \|C\| + \epsilon_B}\right) \tag{124}$$

as claimed. □

Later we will see that certain families of Hamiltonians are extremely powerful simulators: they can simulate any other Hamiltonian.

Definition 25 *We say that a family of Hamiltonians is a* universal simulator*, or is* universal*, if* any *(finite-dimensional) Hamiltonian can be simulated by a Hamiltonian from the family. We say that the universal simulator is* efficient *if the simulation is efficient for all local Hamiltonians.*

Although we restrict to finite-dimensional Hamiltonians in this definition, infinite-dimensional cases can be treated via standard discretisation techniques. Indeed, we will see one such example later. We restrict our notion of efficiency to local Hamiltonians, as this is a natural class of Hamiltonians which have efficient descriptions themselves.

First, however, we will show that the definition of simulation we have arrived at has some interesting consequences.

4.3 Simulation and static properties

First we show that Hamiltonian simulation does indeed approximately preserve important physical properties of the simulated Hamiltonian. Although this is effectively immediate for perfect simulations from the definition of encodings, for approximate simulations we need to check how the level of inaccuracy in the simulation translates into a level of inaccuracy in the property under consideration. We first do this for eigenvalues; essentially the same result was shown in [BH17] but we include a proof for completeness.

Lemma 26 *Let H act on $(\mathbb{C}^d)^{\otimes n}$, let H' be a (Δ, η, ϵ)-simulation of H, and let $\lambda_i(H)$ (resp. $\lambda_i(H')$) be the i'th smallest eigenvalue of H (resp. H'). Then for all $1 \leq i \leq d^n$ and all j such that $(i-1)(p+q)+1 \leq j \leq i(p+q)$, $|\lambda_i(H) - \lambda_j(H')| \leq \epsilon$ (where the integers p, q are those appearing in simulation's encoding).*

Proof For any i, j satisfying the above conditions, $\lambda_i(H) = \lambda_j(\mathcal{E}(H))$ by the definition of an encoding. By condition (i) of Definition 22, the spectrum of $\widetilde{\mathcal{E}}(H)$ is the same as the spectrum of $\mathcal{E}(H)$. By condition (ii) of Definition 22 and Weyl's inequality $|\lambda_j(\widetilde{\mathcal{E}}(H)) - \lambda_j(H')| \leq \|\widetilde{\mathcal{E}}(H) - H'_{\leq\Delta}\|$, each eigenvalue differs from its counterpart by at most ϵ. □

Next we verify that simulation approximately preserves partition functions.

Proposition 27 *Let H' on m d'-dimensional qudits be a (Δ, η, ϵ)-simulation of H on n d-dimensional qudits, with $\|H'_{\leq\Delta} - \widetilde{\mathcal{E}}(H)\| \leq \epsilon$ for some encoding $\widetilde{\mathcal{E}}(H) = \widetilde{V}(H^{\oplus p} \oplus \bar{H}^{\oplus q})\widetilde{V}^\dagger$. Then the relative error in the simulated partition function evaluated at β satisfies*

$$\frac{|\mathcal{Z}_{H'}(\beta) - (p+q)\mathcal{Z}_H(\beta)|}{(p+q)\mathcal{Z}_H(\beta)} \leq \frac{(d')^m e^{-\beta\Delta}}{(p+q)d^n e^{-\beta\|H\|}} + (e^{\epsilon\beta} - 1). \tag{125}$$

Proof Let S be the low-energy subspace of H', $S = \mathrm{Im}(\widetilde{V})$. We have

$$(p+q)\mathcal{Z}_H(\beta) = (p+q)\operatorname{Tr}(e^{-\beta H}) = \operatorname{Tr}(e^{-\beta\widetilde{\mathcal{E}}(H)}|_S) \tag{126}$$

and hence

$$\frac{|\mathcal{Z}_{H'}(\beta) - (p+q)\mathcal{Z}_H(\beta)|}{(p+q)\mathcal{Z}_H(\beta)} \tag{127}$$

$$= \frac{|\operatorname{Tr}(e^{-\beta H'}) - \operatorname{Tr}(e^{-\beta\widetilde{\mathcal{E}}(H)}|_S)|}{\operatorname{Tr}(e^{-\beta\widetilde{\mathcal{E}}(H)}|_S)} \tag{128}$$

$$\leq \frac{|\operatorname{Tr}(e^{-\beta H'}) - \operatorname{Tr}(e^{-\beta H'|_{\leq\Delta}})|}{(p+q)\operatorname{Tr}(e^{-\beta H})} + \frac{|\operatorname{Tr}(e^{-\beta H'|_{\leq\Delta}}) - \operatorname{Tr}(e^{-\beta\widetilde{\mathcal{E}}(H)|_S})|}{\operatorname{Tr}(e^{-\beta\widetilde{\mathcal{E}}(H)}|_S)}. \tag{129}$$

For the first term, the numerator is upper-bounded by $(d')^m e^{-\beta\Delta}$, whereas in the denominator $\operatorname{Tr}(e^{-\beta H})$ is lower-bounded by $d^n e^{-\beta\|H\|}$. For the second term, we write λ_k for the k'th eigenvalue of H (in nonincreasing order), and $\lambda_k + \epsilon_k$ for the k'th eigenvalue of $H'|_{\leq\Delta}$ (in the same order), and have

$$|\operatorname{Tr}(e^{-\beta H'|_{\leq\Delta}}) - \operatorname{Tr}(e^{-\beta\widetilde{\mathcal{E}}(H)|_S})| \leq \sum_k |e^{-\beta(\lambda_k+\epsilon_k)} - e^{-\beta\lambda_k}| = \sum_k e^{-\beta\lambda_k}|e^{-\beta\epsilon_k} - 1|. \tag{130}$$

By Lemma 26, $|\epsilon_k| \leq \epsilon$ for all k, so we have $|e^{-\beta\epsilon_k} - 1| \leq e^{\beta\epsilon} - 1$, and thus the relative error is upper-bounded by

$$\frac{(d')^m e^{-\beta\Delta}}{(p+q)d^n e^{-\beta\|H\|}} + (e^{\epsilon\beta} - 1) \tag{131}$$

as claimed. □

We remark that if we choose $\Delta \gg \|H\| + (m\log d' - n\log d - \log(p+q))/\beta$ and $\epsilon \ll 1/\beta$ then this relative error tends to zero. All the simulations we construct allow us to choose $\Delta \gg m - n$, so these scalings are possible.

4.4 Simulation and time-evolution

We showed in Proposition 4 that encodings allow perfect simulation of time-evolution. We now confirm that this holds for simulations too, up to a small approximation error.

Proposition 28 *Let H' be a (Δ, η, ϵ)-simulation of H with corresponding encoding $\mathcal{E} = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$. Then for any density matrix ρ' in the encoded subspace, so that $\mathcal{E}(\mathbb{1})\rho' = \rho'$,*

$$\|e^{-iH't}\rho' e^{iH't} - e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}\|_1 \leq 2\epsilon t + 4\eta \tag{132}$$

Proof Recall that by the definition of simulation there exists an alternative encoding $\widetilde{\mathcal{E}}(M) = \widetilde{V}(M \otimes P + \bar{M} \otimes Q)\widetilde{V}^\dagger$ such that $\widetilde{\mathcal{E}}(\mathbb{1}) = P_{\leq\Delta(H')}$ and

$\|\widetilde{V} - V\| \leq \eta$. Let $\widetilde{\rho} = \widetilde{V}V^\dagger \rho' V\widetilde{V}^\dagger$. Then

$$\|e^{-iH't}\rho' e^{iH't} - e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}\|_1 \tag{133}$$

$$\leq \|e^{-iH't}\rho' e^{iH't} - e^{-iH't}\widetilde{\rho}e^{iH't}\|_1 + \|e^{-iH't}\widetilde{\rho}e^{iH't} - e^{-i\widetilde{\mathcal{E}}(H)t}\widetilde{\rho}e^{i\widetilde{\mathcal{E}}(H)t}\|_1 + \|e^{-i\widetilde{\mathcal{E}}(H)t}\widetilde{\rho}e^{i\widetilde{\mathcal{E}}(H)t} - e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}\|_1 \tag{134}$$

by the triangle inequality. Since ρ' is in the encoded subspace, we know that $VV^\dagger \rho' VV^\dagger = \rho'$. Therefore Lemma 17 lets us bound the first term by $\|\rho' - \widetilde{\rho}\|_1 \leq 2\|\widetilde{V}V^\dagger - VV^\dagger\| \leq 2\eta$. Similarly, noting that

$$e^{-i\widetilde{\mathcal{E}}(H)t}\widetilde{\rho}e^{i\widetilde{\mathcal{E}}(H)t} = \widetilde{V}V^\dagger e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}V\widetilde{V}^\dagger, \tag{135}$$

we use Lemma 17 to bound the third term by $2\|\widetilde{V}V^\dagger - VV^\dagger\| \leq 2\eta$. Finally, for the second term, we note that $P_{\leq\Delta(H')}\widetilde{\rho} = \widetilde{\rho}$, so $e^{-iH't}\widetilde{\rho}e^{iH't} = e^{-iH'_{\leq\Delta}t}\widetilde{\rho}e^{iH'_{\leq\Delta}t}$, and by Lemma 17 again this term is bounded by

$$2\|e^{iH'_{\leq\Delta}t} - e^{i\widetilde{\mathcal{E}}(H)t}\| \leq 2t\|H'_{\leq\Delta} - \widetilde{\mathcal{E}}(H)\| \leq 2\epsilon t \tag{136}$$

where we have used the matrix inequality $\|e^A - e^B\| \leq \|A-B\|\|e^A\|\|e^{A-B}\|$ [HJ91, Corollary 6.2.32]. □

Corollary 29 *Suppose in addition to the conditions of Corollary 29 that $\mathcal{E}$ is a standard encoding. Let $\mathcal{E}_{\text{state}}(\rho) = V(\rho \otimes \sigma)V^\dagger$ for some state σ satisfying $P\sigma = \sigma$, and let $F(\rho') = \text{Tr}_E[(\mathbb{1} \otimes P)V^\dagger \rho' V]$ as defined in (21). Then*

$$\|e^{-iH't}\mathcal{E}_{\text{state}}(\rho)e^{iH't} - \mathcal{E}_{\text{state}}(e^{-iHt}\rho e^{iHt})\|_1 \leq 2\epsilon t + 4\eta, \tag{137}$$

$$\|F(e^{-iH't}\rho' e^{iH't}) - e^{-iHt}F(\rho')e^{iHt}\|_1 \leq 2\epsilon t + 4\eta. \tag{138}$$

Proof The first statement follows from setting $\rho' = \mathcal{E}_{\text{state}}(\rho)$ in Proposition 28 and noting that $e^{-i\mathcal{E}(H)t}\mathcal{E}_{\text{state}}(\rho)e^{i\mathcal{E}(H)t} = \mathcal{E}_{\text{state}}(e^{-iHt}\rho e^{iHt})$. The second statement follows from $F(e^{-i\mathcal{E}(H)t}\rho' e^{i\mathcal{E}(H)t}) = e^{-iHt}F(\rho')e^{iHt}$ and the fact that F is trace-nonincreasing. □

4.5 Errors and noise

An important question for any simulation technique is how errors affecting the simulator relate to errors on the simulated system. Understanding this in full detail will depend strongly on the physical noise model being considered and the implementation details of the simulation. However, our notion of simulation via local encodings enables us to make some general statements about errors.

First, we show that a local error on the simulator does not map between the forward-evolving and backward-evolving parts of the simulator. This implies the existence of a corresponding local error on the original system by using the F map to extract the forward-evolving part. Second, we show that for the types of encoding used in this paper, a stronger result holds: any local error on an

encoded state is equal to the encoding of a local error on the original system. Finally, we show that, under a reasonable physical assumption, any error on the simulator is close to an error that acts only within the encoded subspace. This allows us to continue to simulate time-evolution and measurement following an error.

Theorem 30 *Let $\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$ be a local encoding, where M acts on n qudits, and let ρ' be a state on the encoded subspace such that $\mathcal{E}(\mathbb{1})\rho' = \rho'$. Let $\mathcal{N}'$ be a CP-map whose Kraus operators each act on at most $l < n$ qudits of the simulator system.*

1. *Let $P' = V(\mathbb{1} \otimes P)V^\dagger$ and $Q' = V(\mathbb{1} \otimes Q)V^\dagger$. Then*

$$P'\mathcal{N}'(\rho') = P'\mathcal{N}'(P'\rho') \quad \text{and} \quad Q'\mathcal{N}'(\rho') = Q'\mathcal{N}'(Q'\rho'). \tag{139}$$

2. *Let $\mathcal{E}_{\text{state}}(\rho) = V(\rho \otimes \sigma)V^\dagger$ for a density matrix σ satisfying $P\sigma = \sigma$. Then the map defined by $\mathcal{N}(\rho) = F(\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho)))$ is a CP-map whose Kraus operators act on at most l qudits of the original system.*

Proof Let $\mathcal{N}'(\rho') = \sum_k N'_k \rho' N'^\dagger_k$. For a given k, the Kraus operator N'_k acts on only l qudits of the simulator system. Therefore N'_k must act trivially on at least one subsystem $\mathcal{H}'_j$. Recall from Theorem 15 that there exists a projector P_{E_j} which acts only on the ancilla E_j such that $(\mathbb{1} \otimes P_{E_j})P = P$ and $(\mathbb{1} \otimes P_{E_j})Q = 0$. Defining $P'_j = \mathbb{1} \otimes V_j P_{E_j} V_j^\dagger$, we have $P'_j P' = P'$ and $P'_j Q' = 0$. Note that P'_j acts non-trivially only on $\mathcal{H}'_j$ and so commutes with N'_k. Therefore

$$P'N'_k\mathcal{E}(\mathbb{1}) = P'P'_jN'_k\mathcal{E}(\mathbb{1}) = P'N'_kP'_j\mathcal{E}(\mathbb{1}) = P'N'_kP'_j(P' + Q') = P'N'_kP'. \tag{140}$$

So, remembering that ρ' is in the encoded subspace and satisfies $\rho' = \mathcal{E}(\mathbb{1})\rho'$, we have

$$P'\mathcal{N}(\rho') = \sum_k P'N'_k\mathcal{E}(\mathbb{1})\rho' N'^\dagger_k = \sum_k P'N'_kP'\rho' N'^\dagger_k = P'\mathcal{N}'(P'\rho'). \tag{141}$$

The statement for Q follows analogously.

We now prove the second part of the theorem. $\mathcal{N}(\rho)$ is clearly CP, since it is defined by a composition of CP maps. Let the spectral decomposition of σ be given by $\sigma = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$. Extend $\{|\psi_j\rangle\}_j$ to a basis for the subspace of the ancilla E given by the support of P. Then

$$\begin{aligned}
\mathcal{N}(\rho) &= F(\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))) && (142)\\
&= \text{Tr}_E[(\mathbb{1} \otimes P)\sum_k V^\dagger N'_k V(\rho \otimes \sigma)V^\dagger N'^\dagger_k V(\mathbb{1} \otimes P)] && (143)\\
&= \sum_{i,j,k} (\mathbb{1} \otimes \langle\psi_i|)V^\dagger N'_k V(\rho \otimes \lambda_j |\psi_j\rangle\langle\psi_j|)V^\dagger N'^\dagger_k V(\mathbb{1} \otimes |\psi_i\rangle) && (144)\\
&= \sum_{k,i,j} N_{i,j,k}\rho N^\dagger_{i,j,k} && (145)
\end{aligned}$$

where $N_{i,j,k} = \sqrt{\lambda_j}(\mathbb{1} \otimes \langle\psi_i|)V^\dagger N'_k V(\mathbb{1} \otimes |\psi_j\rangle)$ are the Kraus operators of $\mathcal{N}$. Since $\mathcal{E}$ is a local encoding, the isometry V may be chosen to be local by Theorem 15, so $V^\dagger N'_k V$ acts non-trivially on at most l qudits of the original system. Therefore the Kraus operators $N_{i,j,k}$ act non-trivially on at most l qudits, as claimed. □

For a general encoding with a corresponding map on states $\mathcal{E}_{\text{state}}(\rho) = V(\rho \otimes \sigma)V^\dagger$, the error $\mathcal{N}'$ may entangle ρ and σ, so it is not possible in general to show that $\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho)) \approx \mathcal{E}_{\text{state}}(\mathcal{N}(\rho))$. However, if $\text{rank}(P) = 1$ (as is the case in all our simulations) then we are able to get a stronger result, which composes more straightforwardly with our other results.

Corollary 31 *Let $\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$ be a local encoding with $\text{rank}(P) = 1$ and let $\mathcal{E}_{\text{state}}(\rho) = V(\rho \otimes P)V^\dagger$. Let $\mathcal{N}'$ and $\mathcal{N}$ be the CP-maps given in Theorem 30. Then*

$$\mathcal{E}(\mathbb{1})\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))\mathcal{E}(\mathbb{1}) = \mathcal{E}_{\text{state}}(\mathcal{N}(\rho)). \tag{146}$$

Proof Let the Kraus operators of $\mathcal{N}'$ be given by N'_k. Since $\text{rank}(P) = 1$, we must have $P = |\psi\rangle\langle\psi|$ for some state $|\psi\rangle$ on the ancilla system E. Since $Q'\mathcal{E}_{\text{state}}(\rho) = 0 = \mathcal{E}_{\text{state}}(\rho)Q'$, where Q' is defined as in Theorem 30, part 1 of that theorem shows that $Q'\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho)) = 0 = \mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))Q'$. Then writing $\mathcal{E}(\mathbb{1}) = P' + Q'$, we have

$$\mathcal{E}(\mathbb{1})\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))\mathcal{E}(\mathbb{1}) = P'\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))P' \tag{147}$$

$$= V(\mathbb{1} \otimes |\psi\rangle\langle\psi|)V^\dagger \left(\sum_k N'_k V(\rho \otimes |\psi\rangle\langle\psi|)V^\dagger N'^\dagger_k\right) V(\mathbb{1} \otimes |\psi\rangle\langle\psi|)V^\dagger \tag{148}$$

$$= V\left(\sum_k N_k \rho N_k^\dagger \otimes |\psi\rangle\langle\psi|\right) V^\dagger = \mathcal{E}_{\text{state}}(\mathcal{N}(\rho)), \tag{149}$$

where we recall from the proof of Theorem 30 that the Kraus operators of $\mathcal{N}(\rho)$ are given by $N_k = (\mathbb{1} \otimes \langle\psi|)V^\dagger N'_k V(\mathbb{1} \otimes |\psi\rangle)$ (the sum over i and j is not necessary when $\text{rank}(P) = 1$). □

Corollary 31 is the strongest general result relating errors on the simulator and simulated systems that one could hope for: it states that any error (CP-map) on the simulator system corresponds naturally to simulating an error (CP-map) on the simulated system.

Even in the more general setting of Theorem 30, we interpret the map $\mathcal{N}(\rho) = F(\mathcal{N}'\mathcal{E}_{\text{state}}(\rho))$ as the error on the original system corresponding to $\mathcal{N}'$. This is because by part 1 of Theorem 30 we have $B(\mathcal{E}_{\text{state}}(\rho)) = 0$, and therefore by (20), for any observable A,

$$\text{Tr}[A\mathcal{N}(\rho)] = \text{Tr}[\mathcal{E}(A)\mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))]. \tag{150}$$

Although $\mathcal{N}'$ may not map between the forwards and backwards parts of the encoded space, it may take a state out of the encoded subspace. But in order

to implement a local measurement with Proposition 13 and time-evolve with Corollary 29, we need $\rho' = \mathcal{N}'(\mathcal{E}_{\text{state}}(\rho))$ to be in the encoded subspace.

The map $\rho' \mapsto \mathcal{E}(\mathbb{1})\mathcal{N}'(\rho')\mathcal{E}(\mathbb{1})$ does map within the encoded subspace, and has the same corresponding error $\mathcal{N}$ on the original system. Indeed, it is the map that appears in Corollary 31. For this error map we can therefore apply Proposition 13 and Corollary 29 as desired. We will make an extra physically-motivated assumption on the form of the error map $\mathcal{N}'$, which guarantees that the difference between this map and $\mathcal{N}'$ is negligible.

Let H' be a (Δ, η, ϵ)-simulation of H with corresponding local encoding $\mathcal{E}$. We might reasonably assume that errors that take the state out of the low-energy space of H' are unlikely due to the high energy required for such an error. We can formalise this by considering only noise operations $\mathcal{N}'$ such that $\mathrm{Tr}[P_{\leq\Delta(H')}\mathcal{N}'(\sigma)] \geq 1 - \delta$ for any state σ supported only on $S_{\leq\Delta(H')}$, and some δ.

Proposition 32 *Let H' be a (Δ, η, ϵ)-simulation of H with corresponding local encoding $\mathcal{E}$. Let $\mathcal{N}'$ be a quantum channel acting on the simulator system and let ρ' be a state in the encoded subspace, so that $\mathcal{E}(\mathbb{1})\rho' = \rho'$.*

Then, if $\mathrm{Tr}[P_{\leq\Delta(H')}\mathcal{N}'(\sigma)] \geq 1-\delta$ for all states σ supported only on $S_{\leq\Delta(H')}$,

$$\|\mathcal{N}'(\rho') - \mathcal{E}(\mathbb{1})\mathcal{N}'(\rho')\mathcal{E}(\mathbb{1})\|_1 \leq \sqrt{\delta(4-3\delta)} + 8\eta. \tag{151}$$

Proof For readability, write $P_{\leq\Delta} := P_{\leq\Delta(H')}$. Then three applications of the triangle inequality give

$$\begin{aligned}
&\|\mathcal{N}'(\rho') - \mathcal{E}(\mathbb{1})\mathcal{N}'(\rho')\mathcal{E}(\mathbb{1})\|_1 && (152)\\
&\quad\leq \|\mathcal{E}(\mathbb{1})\mathcal{N}'(\rho')\mathcal{E}(\mathbb{1}) - \mathcal{E}(\mathbb{1})\mathcal{N}'(\sigma)\mathcal{E}(\mathbb{1})\|_1 && (153)\\
&\qquad + \|\mathcal{E}(\mathbb{1})\mathcal{N}'(\sigma)\mathcal{E}(\mathbb{1}) - P_{\leq\Delta}\mathcal{N}'(\sigma)P_{\leq\Delta}\|_1 && (154)\\
&\qquad + \|P_{\leq\Delta}\mathcal{N}'(\sigma)P_{\leq\Delta} - \mathcal{N}'(\sigma)\|_1 + \|\mathcal{N}'(\sigma) - \mathcal{N}'(\rho')\|_1 && (155)
\end{aligned}$$

where $\sigma = \widetilde{V}V^\dagger\rho' V\widetilde{V}^\dagger$. Since $\mathcal{N}'$ is a quantum channel and $\mathcal{E}(\mathbb{1})$ is a projector, the first and fourth terms are both bounded by

$$\|\rho' - \sigma\|_1 = \|\rho' - \widetilde{V}V^\dagger\rho' V\widetilde{V}^\dagger\|_1 \leq 2\|VV^\dagger - \widetilde{V}V^\dagger\| \leq 2\eta, \tag{156}$$

where we have used Lemma 17. Similarly, we can bound the second term using Lemma 17 twice:

$$\|\mathcal{E}(\mathbb{1})\mathcal{N}'(\sigma)\mathcal{E}(\mathbb{1}) - P_{\leq\Delta}\mathcal{N}'(\sigma)P_{\leq\Delta}\|_1 \leq 2\|\mathcal{E}(\mathbb{1}) - P_{\leq\Delta}\| \leq 4\eta. \tag{157}$$

It remains to bound the third term $\|P_{\leq\Delta}\mathcal{N}'(\sigma)P_{\leq\Delta} - \mathcal{N}'(\sigma)\|$ in terms of δ using the condition assumed in the proposition. Given any state $|\psi\rangle$ such that $P_{\leq\Delta}|\psi\rangle \neq |\psi\rangle$, define the orthonormal states $|\phi_0\rangle = P_{\leq\Delta}|\psi\rangle / \sqrt{1-x}$ and $|\phi_1\rangle = (\mathbb{1} - P_{\leq\Delta})|\psi\rangle / \sqrt{x}$ where $x = 1 - \langle\psi| P_{\leq\Delta}|\psi\rangle$. The operator $|\psi\rangle\langle\psi| - P_{\leq\Delta}|\psi\rangle\langle\psi| P_{\leq\Delta}$ is a rank 2 operator which acts non-trivially only on the space

spanned by $\{|\phi_0\rangle, |\phi_1\rangle\}$ as the following matrix:

$$\begin{pmatrix} 0 & \sqrt{x(1-x)} \\ \sqrt{x(1-x)} & x \end{pmatrix} \quad \text{with eigenvalues } \lambda_\pm = \frac{x}{2} \pm \sqrt{x(1-x) + \frac{x^2}{4}}. \tag{158}$$

Therefore $\||\psi\rangle\langle\psi| - P_{\leq\Delta} |\psi\rangle\langle\psi| P_{\leq\Delta}\|_1 = |\lambda_+| + |\lambda_-| = \sqrt{x(4-3x)}$. This equality also holds trivially in the case $P_{\leq\Delta} |\psi\rangle = |\psi\rangle$.

Using the spectral decomposition, we can write $\mathcal{N}'(\sigma) = \sum_j \lambda_j |\psi_j\rangle\langle\psi_j|$ and use the triangle inequality to show that the third term in (155) is bounded by

$$\|\mathcal{N}'(\sigma) - P_{\leq\Delta}\mathcal{N}'(\sigma)P_{\leq\Delta}\|_1 \tag{159}$$

$$\leq \sum_j \lambda_j \||\psi_j\rangle\langle\psi_j| - P_{\leq\Delta} |\psi_j\rangle\langle\psi_j| P_{\leq\Delta}\|_1 \tag{160}$$

$$= \sum_j \lambda_j \sqrt{x_j(4-3x_j)} = \sum_j \sqrt{\lambda_j x_j}\sqrt{\lambda_j(4-3x_j)} \tag{161}$$

$$\leq \sqrt{\left(\sum_j \lambda_j x_j\right)\left(4 - 3\sum_k \lambda_k x_k)\right)} \tag{162}$$

where $x_j = 1 - \langle\psi_j| P_{\leq\Delta} |\psi_j\rangle$ and we have used the Cauchy-Schwarz inequality in the last step. The result follows from $\sum_j \lambda_j x_j = 1 - \mathrm{Tr}(P_{\leq\Delta}\mathcal{N}'(\sigma)) = \delta$. □

By setting $\rho' = \mathcal{E}_{\mathrm{state}}(\rho)$ in Proposition 32, and using Corollary 31, we have

Corollary 33 *Let H' be a (Δ, η, ϵ)-simulation of H with corresponding local encoding $\mathcal{E}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$ such that* $\mathrm{rank}(P) = 1$. *Let $\mathcal{E}_{\mathrm{state}}(\rho) = V(\rho \otimes P)V^\dagger$ and let $\mathcal{N}'$ be a quantum channel whose Kraus operators act on at most $l < n$ qudits of the simulator system.*

Then, if $\mathrm{Tr}[P_{\leq\Delta(H')}\mathcal{N}'(\mathcal{E}_{\mathrm{state}}(\rho))] \geq 1 - \delta$, there exists a CP-map $\mathcal{N}$ whose Kraus operators act on at most l qudits of the original system such that

$$\|\mathcal{N}'(\mathcal{E}_{\mathrm{state}}(\rho)) - \mathcal{E}_{\mathrm{state}}(\mathcal{N}(\rho))\|_1 \leq \sqrt{\delta(4-3\delta)} + 8\eta. \tag{163}$$

5 Universal Hamiltonian simulation

Having drawn some consequences from the notion of simulation, we will now move on to prove that certain types of Hamiltonians are universal simulators, first introducing the key technique we use: perturbative reductions [KKR06; BH17; OT08; BDL11].

5.1 Techniques

Let $\mathcal{H}_{\mathrm{sim}}$ be a Hilbert space decomposed as $\mathcal{H}_{\mathrm{sim}} = \mathcal{H}_+ \oplus \mathcal{H}_-$, and let $\Pi_\pm$ denote the projector onto $\mathcal{H}_\pm$. For any linear operator O on $\mathcal{H}_{\mathrm{sim}}$, write

$$O_{--} = \Pi_- O \Pi_-, \quad O_{-+} = \Pi_- O \Pi_+, \quad O_{+-} = \Pi_+ O \Pi_-, \quad O_{++} = \Pi_+ O \Pi_+. \tag{164}$$

Let H_0 be a Hamiltonian such that H_0 is block-diagonal with respect to the split $\mathcal{H}_+ \oplus \mathcal{H}_-$, $(H_0)_{--} = 0$, and $\lambda_{\min}((H_0)_{++}) \geq 1$.

Slight variants of the following lemmas were shown in [BH17], building on previous work [OT08; BDL11]:

Lemma 34 (First-order simulation [BH17]) *Let H_0 and H_1 be Hamiltonians acting on the same space. Suppose there exists a local isometry V such that $\mathrm{Im}(V) = \mathcal{H}_-$ and*

$$\|VH_{\mathrm{target}}V^\dagger - (H_1)_{--}\| \leq \epsilon/2. \tag{165}$$

Then $H_{\mathrm{sim}} = \Delta H_0 + H_1$ $(\Delta/2, \eta, \epsilon)$-simulates H_{target}, provided that the bound $\Delta \geq O(\|H_1\|^2/\epsilon + \|H\|_1/\eta)$ holds.

Lemma 35 (Second-order simulation [BH17]) *Let H_0, H_1, H_2 be Hamiltonians acting on the same space, such that: $\max\{\|H_1\|, \|H_2\|\} \leq \Lambda$; H_1 is block-diagonal with respect to the split $\mathcal{H}_+ \oplus \mathcal{H}_-$; and $(H_2)_{--} = 0$. Suppose there exists a local isometry V such that $\mathrm{Im}(V) = \mathcal{H}_-$ and*

$$\|VH_{\mathrm{target}}V^\dagger - (H_1)_{--} + (H_2)_{-+}H_0^{-1}(H_2)_{+-}\| \leq \epsilon/2. \tag{166}$$

Then $H_{\mathrm{sim}} = \Delta H_0 + \Delta^{1/2}H_2 + H_1$ $(\Delta/2, \eta, \epsilon)$-simulates H_{target}, provided that $\Delta \geq O(\Lambda^6/\epsilon^2 + \Lambda^2/\eta^2)$.

Lemma 36 (Third-order simulation [BH17]) *Let H_0, H_1, H_1', H_2 be Hamiltonians acting on the same space, such that: $\max\{\|H_1\|, \|H_1'\|, \|H_2\|\} \leq \Lambda$; H_1 and H_1' are block-diagonal with respect to the split $\mathcal{H}_+ \oplus \mathcal{H}_-$; $(H_2)_{--} = 0$. Suppose there exists a local isometry V such that $\mathrm{Im}(V) = \mathcal{H}_-$ and*

$$\|VH_{\mathrm{target}}V^\dagger - (H_1)_{--} - (H_2)_{-+}H_0^{-1}(H_2)_{++}H_0^{-1}(H_2)_{+-}\| \leq \epsilon/2 \tag{167}$$

and also that

$$(H_1')_{--} = (H_2)_{-+}H_0^{-1}(H_2)_{+-}. \tag{168}$$

Then $H_{\mathrm{sim}} = \Delta H_0 + \Delta^{2/3}H_2 + \Delta^{1/3}H_1' + H_1$ $(\Delta/2, \eta, \epsilon)$-simulates H_{target}, provided that $\Delta \geq O(\Lambda^{12}/\epsilon^3 + \Lambda^3/\eta^3)$.

In fact, whenever we use Lemmas 34 and 35 we will be able to replace the approximate equalities up to $\epsilon/2$ with exact equalities. We do not invoke Lemma 36 explicitly in this work; however, we state it for completeness because it can be used to show that a QMA-completeness result of [OT08] (Theorem 39 below) actually implies a simulation result. The scaling of Δ assumed in these lemmas is sufficient to ensure that $\Delta/2$ separates the high- and low-energy parts of the simulator Hamiltonian H_{sim}. The main difference between these lemmas and their equivalents in [BH17] is that here we insist on locality of the isometry V, corresponding to our local notion of simulation. The correctness proofs of [BH17] go through without change.

We remark that, in order to use the above lemmas, it will often be convenient to add a multiple of the identity to the simulator or target Hamiltonians, corresponding to an overall energy shift. The families of Hamiltonians which we

consider will always contain the identity, so we are free to do this with impunity. For readability, we often omit this implicit freely added identity term when we state the form of restricted types of Hamiltonians below.

In the Hamiltonian complexity literature, many constructions, known as "gadgets", have been developed to prove that special cases of the Local Hamiltonian problem[1] are QMA-complete, by reducing more complex cases to the more specialised cases (e.g. [KKR06; OT08; CM16; PM17]). These reductions often use perturbation theory and can be interpreted as instances of Lemma 34 or Lemma 35. Thus, rather than being merely *reductions*, they are *simulations* in our terminology. Two types of gadget are commonly used:

- **Mediator qubits.** Imagine we have two qubits a and b and would like to implement some effective interaction across them. One way to achieve this is to attach an ancilla, "mediator" qubit c, and apply a heavily-weighted local term H_0 to c, and a less heavily-weighted term $H_2 = H_{ac} + H_{bc}$. If we insist that qubit c is in the ground state of H_0, this produces an effective interaction across qubits a and b, together with some additional local terms on a and b which we can cancel out by adding an extra term H_1. This puts us in the setting of Lemma 35. The isometry V is the map which acts as the identity on qubits a and b, and attaches a qubit c in the ground state of H_0. This type of gadget is used in [OT08; CM16; SV09] and elsewhere in the literature. Whenever such gadgets are used and analysed using second-order perturbation theory, the preconditions of Lemma 35 hold, so we obtain that the physical Hamiltonian constructed simulates the desired logical Hamiltonian.

- **Subspace encodings.** This type of gadget encodes a logical qubit within $\ell = O(1)$ physical qubits. A Hamiltonian H on ℓ qubits is chosen whose ground space is 2-dimensional. Then an overall Hamiltonian is produced using a sum of heavily-weighted H terms, one on each ℓ-tuple of physical qubits. Within the ground space of the whole Hamiltonian, each ℓ-tuple corresponds to a qubit. Less heavily-weighted interactions across ℓ-tuples produce interactions across logical qubits. Lemma 34 and Lemma 35 can be used to show that the simulator Hamiltonian does indeed simulate the target Hamiltonian. Now the isometry V is a tensor product of n isometries, each of which maps a qubit to the ground space of H within the space of ℓ qubits. By choosing the right isometry, corresponding to a choice of basis for this ground space, we obtain desired new interactions across logical qubits.

  This type of gadget is used in [CM16]. However, note that two of the reductions in that work (simulating an arbitrary 2-local qubit Hamiltonian with a Hamiltonian made up of interactions of Heisenberg or XY type) were more complicated. In these reductions H acts on 3 qubits and has a 4-dimensional ground space, corresponding to two logical qubits. Then

[1] The problem of computing the ground-state energy of a k-local Hamiltonian on n qubits, up to $1/\operatorname{poly}(n)$ precision [KSV02; KKR06].

additional heavily weighted terms are used to effectively project one qubit in each logical pair into a fixed, and highly entangled, state of n qubits. This technique would not comply with our notion of simulation, as the state attached by the corresponding isometry V would be far from a product state. Here we no longer need to use this type of reduction as we have a genuinely local simulation (Theorem 40 below).

In this work we will use both of these kinds of simulation. For readability, we will not fully repeat the correctness proofs of the simulations from previous work, instead sketching the arguments and deferring to the original papers for technical details. However, we stress that replacing the analysis of these gadgets in previous work with the use of Lemmas 34 and 35 is sufficient to obtain fully rigorous proofs of correctness.

In addition, to gain some intuition, we now describe more formally how one of the simpler gadgets from [OT08] can be analysed using Lemma 35, and verify that it fits the constraints. The gadget, which is called the subdivision gadget and is an example of a mediator qubit gadget, allows a k-local Hamiltonian to be simulated by a $(\lceil k/2 \rceil + 1)$-local Hamiltonian. Consider an interaction of the form $H_{\text{target}} = A_a B_b$, where A acts on a subset of qubits a, and B acts on a disjoint subset of qubits b. A mediator qubit c is introduced and we take Hamiltonians

$$H_0 = |1\rangle\langle 1|_c, \quad H_2 = \frac{1}{\sqrt{2}}(A_a X_c - X_c B_b). \tag{169}$$

Then $(H_2)_{-+} = (H_2)^\dagger_{+-} = \frac{1}{\sqrt{2}} |0\rangle\langle 1|_c (A_a - B_b)$, so

$$(H_2)_{-+} H_0^{-1} (H_2)_{+-} = \frac{1}{2} |0\rangle\langle 0|_c (A_a - B_b)^2 = |0\rangle\langle 0|_c (\frac{1}{2} A_a^2 - A_a B_b + \frac{1}{2} B_b^2). \tag{170}$$

In addition, $(H_2)_{--} = 0$. We choose $H_1 = \frac{1}{2}(A_a^2 + B_b^2)$, so $(H_1)_{--} = \frac{1}{2} |0\rangle\langle 0|_c (A_a^2 + B_b^2)$. Consider the isometry defined by $V |\psi\rangle_{ab} = |\psi\rangle_{ab} |0\rangle_c$. Then it is easy to verify that

$$V H_{\text{target}} V^\dagger = (H_1)_{--} - (H_2)_{-+} H_0^{-1} (H_2)_{+-}. \tag{171}$$

It follows from Lemma 35 that, for sufficiently high Δ, $H_{\text{sim}} = \Delta H_0 + \sqrt{\Delta} H_2 + H_1$ (Δ, η, ϵ)-simulates H_{target}. Observe that H_{sim} contains interactions on only at most $\max\{|a| + 1, |b| + 1\}$ qubits. This idea can be used to reduce the locality of the whole Hamiltonian simultaneously, by writing each k-local interaction term in the original Hamiltonian as a sum of tensor product interactions, and adding a new mediator qubit for each such interaction to simulate it with a $(\lceil k/2 \rceil + 1)$-local interaction. The corresponding isometry simply attaches a state of $\text{poly}(n)$ qubits, each in the state $|0\rangle$, so is local.

Since each term of H_2 acts on at most one mediator qubit, there is no interference between gadgets and the total effective Hamiltonian is simply the sum of the effective interactions of each gadget. We say that the gadgets are

applied *in parallel*. For a detailed discussion of the parallel application of mediator qubit gadgets, see [PM17]. We formalise this discussion in the following lemma, with the addition of a corresponding result for subspace encoding gadgets.

Lemma 37 *Let the Hamiltonian $H_0 = \sum_i H_0^{(i)}$ be a sum of terms $H_0^{(i)}$ each with ground space energy 0 and acting non-trivially only on disjoint subsets of qudits S_i. Let the ground space projection operator for $H_0^{(i)}$ be given by $P_-^{(i)}$ so the overall ground space projection operator for H_0 is given by $P_- = \prod_i P_-^{(i)}$.*

- *If H_1 can be expressed as a sum of terms, $H_1 = \sum_\alpha H_1^{(\alpha)}$, the first order perturbation satisfies*

$$(H_1)_{--} = \sum_\alpha (H_1^{(\alpha)})_{--}. \tag{172}$$

- ***Mediator gadgets.*** *Let $H_2 = \sum_i H_2^{(i)}$, where $H_2^{(i)}$ acts trivially on all qudits in $\cup_{j\neq i} S_j$. Suppose that all first order terms vanish, i.e. $P_-^{(i)} H_2^{(i)} P_-^{(i)} = 0$ for all i. Then the second and third order terms are given by*

$$-(H_2 H_0^{-1} H_2)_{--} = -\sum_i P_- H_2^{(i)} (H_0^{(i)})^{-1} H_2^{(i)} P_- \tag{173}$$

$$-(H_2 H_0^{-1} H_2 H_0^{-1} H_2)_{--} = -\sum_i P_- H_2^{(i)} (H_0^{(i)})^{-1} H_2^{(i)} (H_0^{(i)})^{-1} H_2^{(i)} P_-. \tag{174}$$

- ***Subspace gadgets.*** *Let $H_2 = \sum_{(i,j)} H_2^{(i,j)}$ for ordered pairs (i,j), where $H_2^{(i,j)}$ acts non-trivially only on S_i and S_j and raises both sets of qudits completely out of their ground spaces such that $P_-^{(i)} H_2^{(i,j)} P_- = 0$ and $P_-^{(j)} H_2^{(i,j)} P_- = 0$. Then the second order perturbation is given by*

$$-(H_2 H_0^{-1} H_2)_{--} = -\sum_{(i,j)} P_- H_2^{(i,j)} \left(H_0^{(i)} + H_0^{(j)} \right)^{-1} H_2^{(i,j)} P_-. \tag{175}$$

Before providing a proof, we remark why different results are needed for mediator and subspace gadgets. In the mediator gadget case, the qudits of S_i are in a one dimensional ground space of $H_0^{(i)}$ for all i, and the effective Hamiltonian acts non-trivially on the remaining qudits in $\overline{\cup_j S_j}$. Therefore interesting interactions can be effected, even when each perturbative term acts on only one of the sets S_i. Whereas for subspace gadgets, the ith logical qudit lives in the groundspace of $H_0^{(i)}$ on the physical qudits S_i, so we need perturbative terms to act between different S_i in order to make 2-local interactions.

Proof The first claim is trivial. For mediator qudit gadgets, define a projection operator $P_-^{(\overline{i})} = \prod_{k\neq i} P_-^{(k)}$ and note that it acts trivially on S_i, so commutes

with $H_2^{(i)}$. Since the ground state energy for each $H_0^{(k)}$ is zero, $H_0^{(k)} P_-^{(k)} = 0$ and so $H_0^{(k)} P_-^{\overline{(i)}} = 0$ for all $k \neq i$. Therefore,

$$(H_0)^{-1} P_-^{\overline{(i)}} = \left(\sum_k H_0^{(k)} \right)^{-1} P_-^{\overline{(i)}} = (H_0^{(i)})^{-1} P_-^{\overline{(i)}}. \tag{176}$$

Since $P_- = P_-^{\overline{(i)}} P_-$, the second order term is given by

$$\begin{aligned} -(H_2 H_0^{-1} H_2)_{--} &= -\sum_i P_- H_2 H_0^{-1} H_2^{(i)} P_- && (177)\\ &= -\sum_i P_- H_2 H_0^{-1} P_-^{\overline{(i)}} H_2^{(i)} P_- && (178)\\ &= -\sum_i P_- H_2 P_-^{\overline{(i)}} (H_0^{(i)})^{-1} H_2^{(i)} P_- && (179)\\ &= -\sum_i P_- H_2^{(i)} (H_0^{(i)})^{-1} H_2^{(i)} P_- && (180) \end{aligned}$$

where the final equality holds because $P_-^{(j)} H_2^{(k)} P_-^{(j)} = 0$ for all $j \neq k$, and so $P_- H_2 P_-^{\overline{(i)}} = P_- H_2^{(i)}$.

Using the same techniques, we can show that the third order term is equal to

$$-\sum_{i,j,k} P_- H_2^{(j)} (H_0^{(j)})^{-1} P_-^{\overline{(j)}} H_2^{(k)} P_-^{\overline{(i)}} (H_0^{(i)})^{-1} H_2^{(i)} P_-. \tag{181}$$

If $k \neq i, j$, then $P_-^{(k)}$ appears in the product expression for both $P_-^{\overline{(i)}}$ and $P_-^{\overline{(j)}}$ and so $P_-^{\overline{(j)}} H_2^{(k)} P_-^{\overline{(i)}} = 0$. We may therefore assume $k = j$ (the proof for $k = i$ proceeds analogously), in which case we have

$$-\sum_{i,j} P_- H_2^{(j)} (H_0^{(j)})^{-1} P_-^{\overline{(j)}} H_2^{(j)} P_-^{\overline{(i)}} (H_0^{(i)})^{-1} H_2^{(i)} P_-. \tag{182}$$

The operator $P_-^{\overline{(j)}}$ commutes with $H_2^{(j)}$ and $P_-^{\overline{(i)}}$, and so, remembering that $P_-^{\overline{(j)}} (H_0^{(i)})^{-1} = 0$ for $i \neq j$, we must have $i = j$, giving the desired result.

The proof is very similar for subspace gadgets, but we instead define a projection operator $P_-^{\overline{(i,j)}} = \prod_{k \neq i,j} P_-^{(k)}$ for ordered pairs (i, j), noting that it acts trivially on S_i and S_j, so commutes with $H_2^{(i,j)}$. As before, we have $H_0^{(k)} P_-^{\overline{(i,j)}} = 0$ for all $k \neq i, j$, so $(H_0)^{-1} P_-^{\overline{(i,j)}} = (H_0^{(i)} + H_0^{(j)})^{-1} P_-^{\overline{(i,j)}}$. Therefore

the second order term is given by

$$-(H_2 H_0^{-1} H_2)_{--} = -\sum_{(i,j)} P_- H_2 H_0^{-1} H_2^{(i,j)} P_- \tag{183}$$

$$= -\sum_{(i,j)} P_- H_2 H_0^{-1} P_-^{\overline{(i,j)}} H_2^{(i,j)} P_- \tag{184}$$

$$= -\sum_{(i,j)} P_- H_2 P_-^{\overline{(i,j)}} \left(H_0^{(i)} + H_0^{(j)}\right)^{-1} H_2^{(i,j)} P_- \tag{185}$$

$$= -\sum_{(i,j)} P_- H_2^{(i,j)} \left(H_0^{(i)} + H_0^{(j)}\right)^{-1} H_2^{(i,j)} P_- \tag{186}$$

where the final equality holds since by the form of $H_2^{(i,j)}$ assumed in the lemma, $P_- H_2^{(i',j')} P_-^{\overline{(i,j)}} = 0$ unless $(i',j') = (i,j)$, so $P_- H_2 P_-^{\overline{(i,j)}} = P_- H_2^{(i,j)} P_-^{\overline{(i,j)}}$. □

5.2 Universal simulators

We are now ready to prove universality of a variety of classes of Hamiltonians. The overall structure of the argument is illustrated in Figure 3; the eventual result is that each of the classes of qudit Hamiltonians illustrated in the diagram is universal. For brevity, when we state and prove simulation results, rather than writing "The family of A-Hamiltonians can simulate the family of B-Hamiltonians" for some A and B, we simply write "A-Hamiltonians can simulate B-Hamiltonians". We stress that such a statement is nevertheless rigorous and should be understood in the sense of Definition 22.

We have already proven some of the simulation results required (Lemmas 20 and 21). We now complete the programme of Figure 3 by showing that every remaining type of qudit Hamiltonian in the diagram is universal. The simulation of Lemma 21 may produce terms which include even numbers of Y components. First we show that such terms are not necessary. Note that it was already known that Hamiltonians without any Y components can be QMA-complete [BL08]; what we show here is that such Hamiltonians can in fact be universal simulators.

Lemma 38 *Real k-local qubit Hamiltonians can be simulated by real $(k+1)$-local qubit Hamiltonians whose Pauli decomposition does not contain any Y terms.*

Proof Let H be a real k-local qubit Hamiltonian. For each k'-local interaction h in the Pauli decomposition of H ($k' \leq k$), add an additional mediator qubit a. Since H is real, there must be an even number of Y terms in h. We may assume, by reordering qubits if necessary, that $h = Y^{\otimes 2m} \otimes A$ where A is a tensor product of X and Z terms on $k' - 2m$ qubits.

We use second-order perturbation theory (Lemma 34) to effectively generate h from an interaction containing only X and Z terms. Consider a heavy interaction term H_0 acting only on the mediator qubit, $H_0 = (\mathbb{1} + Z_a)/2 = |0\rangle\langle 0|_a$, with groundstate $|1\rangle_a$, and a perturbative term $H_2 = X_a(X^{\otimes 2m} \otimes \mathbb{1} + (-1)^{m+1} Z^{\otimes 2m} \otimes$

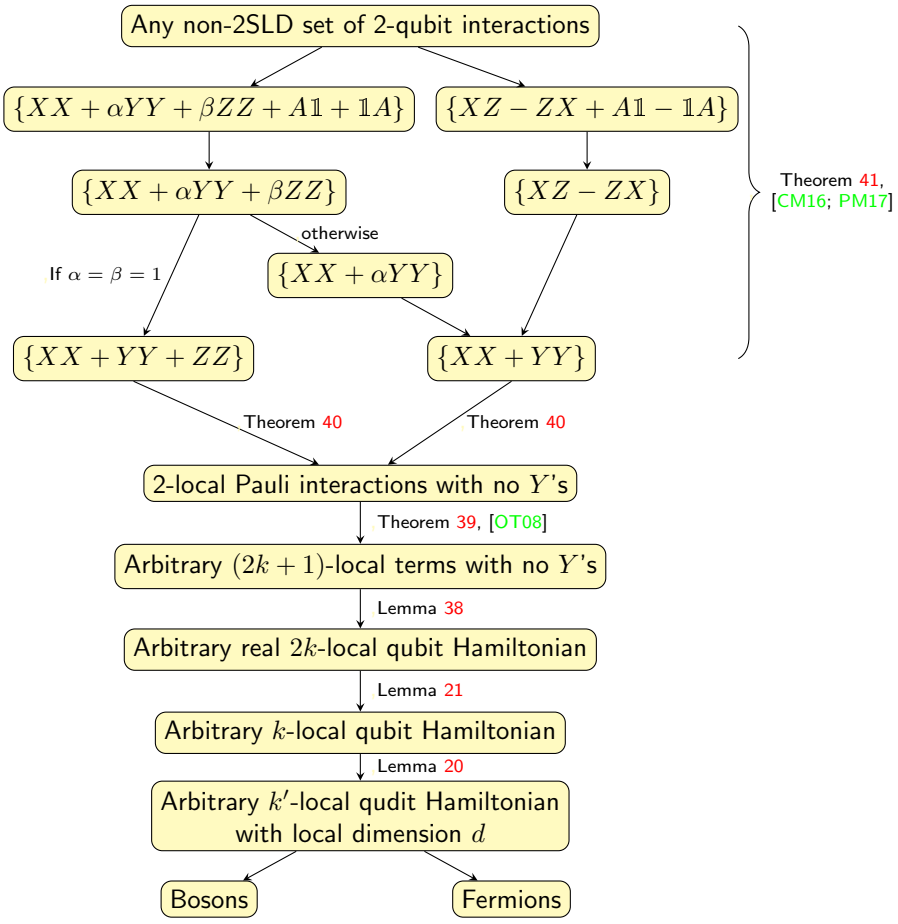


Figure 3: Sequence of simulations used in this work. An arrow from one box to another indicates that a Hamiltonian of the first type can simulate a Hamiltonian of the second type. Where two arrows leave a box, this indicates that a Hamiltonian of this type can simulate one of the two target Hamiltonians, but not necessarily both. "2SLD" is short for "the 2-local parts of all interactions in the set are simultaneously locally diagonalisable", and $k, k' \geq 2$ are arbitrary integers such that $k \geq \lceil k' \log_2 d \rceil$.

A). H_2 acts as a switch between the ground space and the excited space. It is clear that the first-order term $\Pi_- H_2 \Pi_-$ vanishes. The second-order term is, up to a multiple of the identity, of the desired form:

$$-\Pi_- H_2 (H_0^{-1})_{++} H_2 \Pi_- = -|0\rangle\langle 0|_a \left(X^{\otimes 2m} \otimes \mathbb{1} + (-1)^{m+1} Z^{\otimes 2m} \otimes A\right)^2 \quad (187)$$

$$= 2|0\rangle\langle 0|_a \left(Y^{\otimes 2m} \otimes A + \mathbb{1}\right). \quad (188)$$

It follows from Lemma 34 that, for sufficiently large Δ, $H' = \Delta H_0 + \Delta^{1/2} H_2$ is a (Δ, η, ϵ)-simulation of the interaction h. This can be used to simulate the whole Hamiltonian H by applying separate mediator qubit gadgets for each term h in parallel; by Lemma 37, different terms do not interfere with each other. □

It may be tempting to think that a similar second-order mediator qubit gadget could be used to simulate a 1-local Y interaction, since $ZX = iY$. However the same trick would not work if we took $H_2 = X_a(X_1 + Z_1)$, for example, because the anticommutator $\{X, Z\}$ vanishes and so $(X + Z)^2 = 2\mathbb{1}$. Of course, this should not be surprising, as the perturbative expansion of any real Hamiltonian can only result in real Hamiltonian terms.

Next we use a result of Oliveira and Terhal [OT08] to further specialise the class of Hamiltonians proven universal in Lemma 38.

Theorem 39 (essentially [OT08]) *k-local qubit Hamiltonians whose Pauli decomposition does not contain any Y terms can be simulated by 2-local Hamiltonians of the form $\sum_{i>j} \alpha_{ij} A_{ij} + \sum_k (\beta_k X_k + \gamma_k Z_k)$, where A_{ij} is one of the interactions $X_i X_j, X_i Z_j, Z_i X_j$ or $Z_i Z_j$ and $\alpha_{ij}, \beta_k, \gamma_k \in \mathbb{R}$.*

We sketch the proof of Theorem 39; see [OT08] for more technical details.

Proof (sketch) The claim is trivial for $k \leq 2$, so assume $k \geq 3$. We first note that, for each k-tuple of qubits, one can decompose any interaction across that k-tuple as a weighted sum of interactions which are each tensor products of Pauli matrices. These can be thought of as separate hyperedges in the hypergraph of interactions in H, and henceforth treated separately.

Then, to effectively produce each of these Pauli interactions, the subdivision gadgets described in [OT08] can be used. There are two of these gadgets. One gadget simulates an arbitrary k-wise interaction of the form $A \otimes B$ across sets of qubits a and b by using a mediator qubit c, and $\lceil k/2 \rceil$-wise interactions of the form $A_a X_c + X_c B_b$. This gadget was discussed in detail near the start of Section 5.1. Repeated use of this procedure enables k-local interactions to be simulated using 3-local interactions. The second gadget simulates a 3-local Hamiltonian with a 2-local Hamiltonian. The gadget generates interactions of the form $A_a B_b C_c$ by introducing a mediator qubit d and a Hamiltonian whose terms are proportional to $A_a X_d$, $B_b X_d$ and $C_c |1\rangle\langle 1|_d$, and using third-order perturbation theory to generate effective 3-local terms from these [BH17; OT08]. This leads to unwanted 2-local and 1-local terms being generated too, which can be effectively deleted using compensating terms of the form XZ, X, Z. By Lemma 37, these third order mediator qubit gadgets do not interfere. Note that

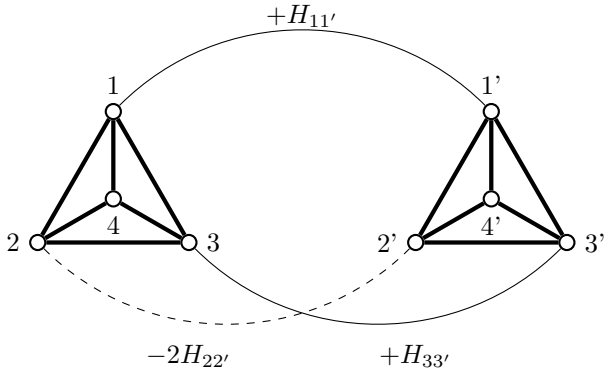


Figure 4: One logical qubit is encoded within a quadruple of physical qubits (1–4 and $1'$–$4'$). 2-local interactions are implemented using interactions across the quadruples. The figure illustrates the Hamiltonian for simulating $X_L X_L$, up to 1-local terms.

the analysis of [OT08] can be replaced with the use of Lemma 36 to show that this gadget indeed gives a simulation in our terminology.

Finally, observe that these gadgets do not introduce any Y terms if they were not present already. □

Next we show that the Heisenberg and XY interactions are sufficient to simulate any Hamiltonian of the form of Theorem 39. This is the most technically involved simulation in this paper. Unlike the argument used in [CM16], here the encoding used is local.

Theorem 40 *Let $\mathcal{F}$ be the family of qubit Hamiltonians of the form $H = \sum_{i>j} \alpha_{ij} A_{ij} + \sum_k (\beta_k X_k + \gamma_k Z_k)$, where A_{ij} is one of the interactions X_iX_j, X_iZ_j, Z_iX_j or Z_iZ_j and $\alpha_{ij}, \beta_k, \gamma_k \in \mathbb{R}$. Then $\{XX+YY+ZZ\}$-Hamiltonians and $\{XX+YY\}$-Hamiltonians can simulate $\mathcal{F}$.*

Proof We prove the claim for the Heisenberg interaction $XX + YY + ZZ$; the argument is completely analogous for the XY interaction $XX + YY$. We use a subspace encoding gadget to encode a logical qubit in the ground space of the Hamiltonian of the complete graph on 4 qubits, as illustrated in Figure 4.

The overall heavy interaction used is

$$H_0 := H_{12} + H_{23} + H_{34} + H_{14} + H_{24} + H_{13} + 6\mathbb{1}, \tag{189}$$

where we write $H_{ij} = X_iX_j + Y_iY_j + Z_iZ_j$. The identity term is present to ensure that the ground space of H_0 corresponds to eigenvalue zero. H_0 has a two dimensional ground space S given in terms of singlet states $|\Psi^-\rangle$ by

$$S = \operatorname{span}\left\{|\Psi^-\rangle_{12}|\Psi^-\rangle_{34}, |\Psi^-\rangle_{13}|\Psi^-\rangle_{24}\right\} \quad \text{where } |\Psi^-\rangle = \frac{|01\rangle - |10\rangle}{\sqrt{2}}. \tag{190}$$

(i,j)	$\Pi_- X_i X_j \Pi_-$
$(1,3)$ $(2,4)$	$-\frac{2}{3}Z_L - \frac{1}{3}\mathbb{1}$
$(1,2)$ $(3,4)$	$-\frac{1}{\sqrt{3}}X_L + \frac{1}{3}Z_L - \frac{1}{3}\mathbb{1}$
$(1,4)$ $(2,3)$	$\frac{1}{\sqrt{3}}X_L + \frac{1}{3}Z_L - \frac{1}{3}\mathbb{1}$

(193)

Table 1: Effective interactions produced by physical interaction acting on different choices of qubits.

We choose the following orthonormal basis for our logical qubit:

$$|0_L\rangle = |\Psi^-\rangle_{13}|\Psi^-\rangle_{24} \qquad |1_L\rangle = \tfrac{2}{\sqrt{3}}|\Psi^-\rangle_{12}|\Psi^-\rangle_{34} - \tfrac{1}{\sqrt{3}}|\Psi^-\rangle_{13}|\Psi^-\rangle_{24} \quad (191)$$

First-order perturbations We can simulate 1-local interactions X_L and Z_L using first-order perturbation theory. By Lemma 34, given a perturbation term H_1, the first-order perturbation is given by $\Pi_- H_1 \Pi_-$, where Π_- is the projector into the ground space. Note that the ground space is defined in terms of singlet states which have the same form in any local basis, and so

$$\Pi_- X_i X_j \Pi_- = \Pi_- Y_i Y_j \Pi_- = \Pi_- Z_i Z_j \Pi_- \quad (192)$$

which we can also check explicitly. Although the heavy Hamiltonian H_0 is invariant under permutations of the physical qubits, this symmetry is lost when we fix the basis, and so $\Pi_- X_i X_j \Pi_-$ does depend on (i,j) – the values are given in Table 1.

Therefore, we can simulate any real 1-local interaction up to an irrelevant identity term; by Lemma 34, choosing $H_1 = \frac{\alpha}{\sqrt{3}}H_{14} + \frac{1}{2}(\frac{\alpha}{\sqrt{3}} - \beta)H_{13}$ will simulate the interaction $\Pi_- H_1 \Pi_- = \alpha X_L + \beta Z_L + \frac{1}{2}(\beta - \sqrt{3}\alpha)\mathbb{1}$.

Second-order perturbations In order to make an effective interaction between two logical qubits we need to use physical interactions that act between two of these 4-qubit gadgets. We label the four physical qubits of one logical qubit as $1,2,3,4$, and the qubits of a second logical qubit with a dash $1',2',3',4'$ and consider a perturbation term of the form $H_2 = \sum \alpha_{ij} H_{ij'}$. All first-order perturbation terms vanish as it is easy to show that $\Pi_- X_i \Pi_- = \Pi_- Y_i \Pi_- = \Pi_- Z_i \Pi_- = 0$ for all $i \in \{1,2,3,4\}$.

Let $H_0^{\text{tot}} = H_0 \otimes \mathbb{1} + \mathbb{1} \otimes H_0$ be the total heavy Hamiltonian on these 8 qubits, and let Π_-^{tot} project onto the groundspace of H_0^{tot}.

We note that $Z_1|\Psi^-\rangle_{12}|\Psi^-\rangle_{34} = |\Psi^+\rangle_{12}|\Psi^-\rangle_{34}$ is an eigenvector of H_0 with eigenvalue 4, where $|\Psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$. Since the other eigenvector spanning the ground space of H_0, $|\Psi^-\rangle_{12}|\Psi^-\rangle_{34}$, is of a similar form, it is clear that Z_1 maps the ground space of H_0 into the eigenspace of eigenvalue 4. By

H_2	2-local part of effective interaction
$H_{11'} \mp H_{33'}$	$\pm Z_L Z_L$
$H_{13'} - H_{11'} \pm H_{32'}$	$\pm Z_L X_L$
$H_{11'} - 2H_{22'} + H_{33'}$	$X_L X_L$
$35H_{11'} + 5H_{22'} - 3H_{33'} + 5H_{44'}$	$-X_L X_L$

Table 2: Effective 2-local interactions produced from different choices of H_2 terms, up to a non-negative scaling factor.

unitary invariance of the Heisenberg interaction, and the symmetry between qubits 1, 2, 3, 4, we can say the same for any X_i, Y_i or Z_i. This allows us to simplify the calculation of the second-order perturbation term,

$$- \Pi_-^{\mathrm{tot}} H_2 \Pi_+ (H_0^{\mathrm{tot}})^{-1} \Pi_+ H_2 \Pi_-^{\mathrm{tot}} \tag{194}$$

$$= -\Pi_-^{\mathrm{tot}} H_2 \Pi_+ \tfrac{1}{4+4} \Pi_+ H_2 \Pi_-^{\mathrm{tot}} = -\tfrac{1}{8} \Pi_-^{\mathrm{tot}} H_2^2 \Pi_-^{\mathrm{tot}} \tag{195}$$

$$= -\tfrac{1}{8} (\Pi_- \otimes \Pi_-) \left(\sum_{i,j,k,l=1}^{4} \alpha_{ij} \alpha_{kl} H_{ij'} H_{kl'} \right) (\Pi_- \otimes \Pi_-) \tag{196}$$

$$= -\tfrac{1}{8} \sum_{i,j,k,l=1}^{4} \alpha_{ij} \alpha_{kl} \big[(\Pi_- X_i X_k \Pi_-) \otimes (\Pi_- X_{j'} X_{l'} \Pi_-) \tag{197}$$

$$+ (\Pi_- X_i Y_k \Pi_-) \otimes (\Pi_- X_{j'} Y_{l'} \Pi_-) + \dots \big]. \tag{198}$$

Next, one can check that $\Pi_- X_i Y_k \Pi_- = \Pi_- X_i Z_k \Pi_- = \Pi_- Y_i Z_k \Pi_- = 0$ for any pair (i, k), so many of these terms vanish. Remembering also that $\Pi_- X_i X_k \Pi_- = \Pi_- Y_i Y_k \Pi_- = \Pi_- Z_i Z_k \Pi_-$, this expression simplifies to

$$-\tfrac{1}{8} \Pi_- H_2^2 \Pi_- = -\tfrac{1}{8} \sum_{i,j,k,l=1}^{4} 3 \alpha_{ij} \alpha_{kl} (\Pi_- X_i X_k \Pi_-) \otimes (\Pi_- X_{j'} X_{l'} \Pi_-), \tag{199}$$

where the effective interactions produced by $\Pi_- X_i X_k \Pi_-$ can be read off again from Table 1.

By Lemma 35, for any $\epsilon > 0$ and sufficiently large $\Delta = \mathrm{poly}(\|H\|, 1/\eta, 1/\epsilon)$, $\Delta H_0 + \Delta^{\frac{1}{2}} H_2 + H_1$ (Δ, η, ϵ)-simulates the interaction $\Pi_- H_1 \Pi_- - \frac{1}{8} \Pi_-^{\mathrm{tot}} H_2^2 \Pi_-^{\mathrm{tot}}$. Choosing H_1 as above we can cancel out any 1-local part of $\frac{1}{8} \Pi_-^{\mathrm{tot}} H_2^2 \Pi_-^{\mathrm{tot}}$, so we are interested only in the 2-local part. Table 2 shows some choices of H_2 with integer coefficients that generate effective interactions whose 2-local part is proportional to $\pm ZZ$, $\pm ZX$, $\pm XX$.

By Lemma 37, we can apply this simulation to each interaction in H in parallel. Letting H' denote the overall simulator Hamiltonian, we finally obtain that, for any $\epsilon > 0$ and sufficiently large $\Delta = \mathrm{poly}(\|H\|, 1/\eta, 1/\epsilon)$, H' is a (Δ, η, ϵ)-simulation of H.

Simulator interaction H	Simulated interaction H'	Gadget
$XX + \alpha YY$	$XX + YY$	$H_{ab} + H_{bc}$
$XX + \alpha YY + \beta ZZ$	$XX + \alpha' YY$	$H_{ab} - H_{bc}$
$XZ - ZX$	$XX + YY$	$H_{ab} + H_{bc} + H_{ca}$

Table 3: Subspace encodings used in Theorem 41. In each case a qubit is encoded within the ground space of H acting on three qubits labelled a–c. Here α, β, α' are fixed nonzero real numbers.

Everything follows through in exactly the same way for the XY interaction. If we set $H_{ij} = X_iX_j + Y_iY_j$ and use the same gadget, the ground space is left unchanged. So the only thing to check is that X_i, Y_i, Z_i all map the ground space into an eigenspace of H_0 again (which they do!). Then the simulated interactions will be the same up to a constant factor of $2/3$. □

Finally, we show that every remaining class of qudit Hamiltonians in Figure 3 can simulate either XY interactions or Heisenberg interactions, implying that they are all universal too.

Theorem 41 *Let $\mathcal{S}$ be a set of interactions on at most 2 qubits. Assume that there does not exist $U \in SU(2)$ such that, for each 2-qubit matrix $H_i \in \mathcal{S}$, $U^{\otimes 2} H_i (U^\dagger)^{\otimes 2} = \alpha_i Z^{\otimes 2} + A_i \otimes \mathbb{1} + \mathbb{1} \otimes B_i$, where $\alpha_i \in \mathbb{R}$ and A_i, B_i are arbitrary single-qubit Hamiltonians. Then $\mathcal{S}$-Hamiltonians can simulate either $\{XX + YY + ZZ\}$-Hamiltonians or $\{XX + YY\}$-Hamiltonians. Furthermore, if the interaction graph of the target Hamiltonian is a 2D square lattice, then the simulator Hamiltonian may also be chosen to be on a 2D square lattice.*

Observe that the assumption in the theorem is equivalent to assuming that the set formed by extracting the 2-local parts of each interaction in $\mathcal{S}$ is not simultaneously locally diagonalisable. Theorem 41 was first proven in [CM16], with the restriction to 2D square lattices shown in [PM17]. These proofs use different terminology (e.g. they prove "reductions" rather than "simulations"). However, all the gadgets used are examples of mediator qubit gadgets or first order subspace encoding gadgets which, as described in Section 5.1, give simulations in our terminology. We therefore restrict ourselves here to sketching the arguments of [CM16; PM17]. See [CM16; PM17] for a full proof of correctness and technical details.

Proof (sketch) The claim follows by chaining together various simulations from [CM16] in the same order as used in that work; the sequence of simulations used is illustrated in Figure 3. To prove the final part of the theorem, each of the gadgets used in [CM16] can be replaced with a gadget from [PM17] which fits onto a square lattice. Most of the steps of the argument show that, given access to one interaction H, we can effectively produce another interaction H'. Three of these are listed in Table 3. These are all simulations of the subspace encoding type, where we encode one logical qubit within the ground

space of a Hamiltonian on 3 physical qubits. The simulations can be analysed using Lemma 34 and, as they are subspace encodings, satisfy the definition of simulation. By applying the right interactions across qubit triples, we obtain new effective interactions between logical qubits. The effective interactions produced are calculated in [CM16]. Alternatively the mediator qubit gadgets of Figure 8 and Figure 11 of [PM17] may be used to perform the same simulations on a square lattice.

A somewhat different case is the interaction $H = XX + \alpha YY + \beta ZZ + A\mathbb{1} + \mathbb{1}A$, where at least one of α and β is nonzero. Here the available interaction corresponds to one which was considered in Table 3, but with an additional 1-local term of some form. The simulation deletes these 1-local terms by introducing 4 ancilla qubits for each logical qubit a. Labelling these qubits a–d, it turns out that the ground state of $H_0 = H_{ab} + H_{cd} - H_{ac} - H_{bd}$ is unique and maximally-entangled across the $(a - c : d)$ split. If these four qubits are forced to be in this state, applying a $-H$ interaction between 4 and a corresponds to a $-A$ term applied to a. This allows the local A terms to be effectively deleted for each H interaction used. The corresponding isometry V attaches 4 ancilla qubits for each of the original qubits, in the ground state of H_0. The interaction $H = XZ - ZX + A\mathbb{1} - \mathbb{1}A$ is similar; here the local part of H can be deleted using $H_0 = H_{ab} + H_{bc} + H_{cd} + H_{da}$. Section 4.6 of [PM17] shows how these gadget constructions may be adjusted slightly such that they fit onto a 2D square lattice.

Now that these special cases have been dealt with, to complete the argument we need to consider an arbitrary set $\mathcal{S}$ of 2-qubit interactions where there is no $U \in SU(2)$ such that, for each 2-qubit matrix $H_i \in \mathcal{S}$, $U^{\otimes 2} H_i (U^\dagger)^{\otimes 2} = \alpha_i Z^{\otimes 2} + A_i \mathbb{1} + \mathbb{1} B_i$. We sketch the argument and defer to [CM16] for details.

Any 2-qubit interaction H_i can be decomposed in terms of parts which are symmetric and antisymmetric under interchange of the qubits on which it acts, and each of these parts can be extracted by taking linear combinations of H_i and the interaction obtained by swapping the two qubits; so we can assume that all the interactions in $\mathcal{S}$ are either symmetric or antisymmetric. The 2-local part of any symmetric interaction H_i can be written as $\sum_{s,t \in \{x,y,z\}} M^{(i)}_{st} \sigma_s \otimes \sigma_t$ for some symmetric 3×3 matrix $M^{(i)}$. Define the Pauli rank of H_i to be the rank of $M^{(i)}$. If there exists $H_i \in \mathcal{S}$ with Pauli rank 2, we consider Hamiltonians produced only using H_i interactions. As discussed in Section 4, by applying local unitaries and up to rescaling and relabelling Pauli matrices, we can replace H_i with $XX + \alpha YY + \beta ZZ + A\mathbb{1} + \mathbb{1}A$ for some A, and some $\alpha, \beta \in \mathbb{R}$ such that at least one of them is nonzero. This is the special case we just considered.

Otherwise, all $H_i \in \mathcal{S}$ have Pauli rank 1; we also know that there must exist $H_i, H_j \in \mathcal{S}$ such that the 2-local parts of H_i and H_j do not commute, by the assumptions of the theorem. This implies that there must exist some linear combination of H_i and H_j which has Pauli rank at least 2. Considering this linear combination, we are back in the same special case as before. Finally, the case where $\mathcal{S}$ contains an antisymmetric interaction can be dealt with in a similar way, by using local unitaries to put that interaction into the previously

considered canonical form $XZ - ZX + A\mathbb{1} - \mathbb{1}A$. □

We finally observe that it was shown in [PM17] that certain interactions remain universal even if they are only permitted to occur with non-negative weights. Indeed, that work showed that the class of qubit Hamiltonians whose interactions are of the form $\alpha XX + \beta YY + \gamma ZZ$, where $\{\alpha+\beta, \alpha+\gamma, \beta+\gamma\} > 0$, can simulate qubit Hamiltonians with arbitrarily positively or negatively weighted interactions of the form $\alpha' XX + \beta' YY + \gamma' ZZ$, for some α', β', γ' such that at least two of α', β', γ' are nonzero. This implies, for example, that the antiferromagnetic Heisenberg interaction is universal.

5.3 Indistinguishable particles

Throughout this work so far, we have only considered Hamiltonians on distinguishable particles with finite-dimensional Hilbert spaces. As stated, our results – and even the definitions of Hamiltonian encoding and simulation – do not apply to indistinguishable particles or infinite-dimensional Hilbert spaces. Extending these definitions to arbitrary self-adjoint operators on infinite-dimensional Hilbert spaces is beyond the scope of the present article.[1]

However, as bosonic and fermionic systems are ubiquitous in many-body physics, and our main focus is to show that there exist simple, universal quantum models that are able to simulate the physics of any other physical system, we will address the question of whether universal spin models such as the Heisenberg- and XY-models can simulate indistinguishable particles. In fact, the required simulations follow from standard techniques for mapping fermionic and bosonic operators to spin operators, so we only sketch the arguments here.

5.3.1 Fermions

The canonical anti-commutation relation (CAR) algebra describing fermions is generated by fermionic creation and annihilation operators c_i, $c_i^\dagger$ satisfying $\{c_i, c_j\} = 0$ and $\{c_i, c_j^\dagger\} = \delta_{ij}$ (where the subscript indexes different fermionic modes). This algebra is finite-dimensional (as long as the single-particle Hilbert space is). It is well known that this algebra can be embedded into an operator algebra acting on a many-qubit system, e.g. by the well-known Jordan-Wigner transformation: $c_i = -\bigotimes_{j<i} Z_i \otimes \frac{X_i + iY_i}{2}$, where we define some arbitrary total-ordering on the qubits. However, this is not sufficient for our purposes. It transforms individual fermionic creation or annihilation operators into operators that act non-trivially on all qubits in the system, so does not give a local encoding.

The mapping introduced by Bravyi and Kitaev [BK02] improves this to $\log n$-local operators (where n is the total number of qubits)[2]. However, simulating these $\log n$-local interactions using a universal model with two-body interactions,

[1] As the definition and characterisation of encodings in particular is very C^*-algebraic in character, it does not seem too difficult to generalise.

[2] Very recent independent work has given an analysis and comparison of different fermion-to-qubit mappings [HTW17].

such as the Heisenberg- or XY-model, will require local interactions whose norms scale super-polynomially in n. Whilst this gives a simulation with polynomial overhead in terms of the system size, it is not strictly speaking efficient according to our definition due to this super-polynomial scaling of the local interaction strengths.

Both of these mappings produce qubit Hamiltonians with the same number of qubits as fermionic modes. This is much stronger than required for an efficient simulation in the the spirit of Definition 22, which allows a polynomial overhead in the simulator system size. The fermion-to-spin mappings studied in [VC05; Bal05; FS14; WHT16] preserve locality by adding additional auxiliary fermionic modes before mapping to qubits, at the expense of a polynomial system-size overhead. The auxiliary fermions must be restricted to the appropriate subspace, which can be done by adding strong local terms to the Hamiltonian (see [Bal05; VC05]). (These strong local terms mutually commute, and when transformed to spin operators become products of Paulis. So these terms in fact form a stabilizer Hamiltonian.) Together with these strong local terms, this mapping gives a spin Hamiltonian that exactly reproduces the original fermionic Hamiltonian in its low-energy subspace. The resulting spin Hamiltonian is local if the simulated fermionic system is a regular lattice Hamiltonian containing only even products of fermionic creation and annihilation operators [VC05]. Simulating the resulting spin Hamiltonian using any universal model then gives an efficient simulation for this important class of fermionic Hamiltonians.

5.3.2 Bosons

In the case of bosons, the canonical commutation relation (CCR) algebra, generated by bosonic creation and annihilation operators a_i, $a_i^\dagger$ satisfying $[a_i, a_j] = 0$ and $[a_i, a_j^\dagger] = \delta_{ij}$, is infinite-dimensional. To simulate bosons with spins, one must necessarily restrict to some finite-dimensional subspace of the full Hilbert space, and only simulate the system within that subspace. The appropriate choice of subspace will depend on the particular bosonic system, and which physics one wishes to simulate, so one cannot give a completely general result here.

However, a natural choice will often be to limit the maximum number of bosons to some finite value N, i.e. to restrict to the finite-dimensional subspace spanned by eigenstates of the total number operator $\sum_i a_i^\dagger a_i$ with eigenvalue $\leq N$. For systems containing multiple bosonic modes, we can alternatively limit the maximum number of bosons in each mode separately, i.e. restrict to the subspace spanned by eigenvectors with eigenvalue $\leq N$ for each $a_i^\dagger a_i$ individually. (Since $[a_i^\dagger a_i, a_j^\dagger a_j] = 0$, this subspace also contains the subspace with maximum *total* number of bosons N.)

In this way, each bosonic mode is restricted individually to a finite-dimensional subspace that can be represented by the Hilbert space of a qudit. The original bosonic Hamiltonian restricted to this subspace is clearly equivalent to some Hamiltonian on these qudits. Furthermore, since $[a_i, a_j] = [a_i, a_j^\dagger] = [a_i^\dagger, a_j^\dagger] = 0$ for $i \neq j$, k-particle bosonic interactions become k-local interactions on the

qudits. The resulting k-local qudit Hamiltonian can then be simulated by the universal model, as shown in previous sections.

In fact, restricting the bosonic creation and annihilation operators to the finite-particle-number subspace in this way is a well-known procedure. The equivalent qudit operators $S_i^{\pm}$ are given by the (exact) Holstein-Primakov transformation [HP40]:

$$S_i^+ = \sqrt{d-1}\sqrt{1 - \frac{a_i^\dagger a_i}{d-1}}\, a_i, \qquad S_i^- = \sqrt{d-1}\, a_i^\dagger \sqrt{1 - \frac{a_i^\dagger a_i}{d-1}}. \tag{200}$$

5.4 Universal stoquastic simulators

It was previously shown by Bravyi and Hastings [BH17] that the Ising model with transverse fields acts as a universal simulator for the class of stoquastic 2-local Hamiltonians. The transverse Ising model (TIM) corresponds to Hamiltonians which can be written as a weighted sum of terms picked from the set $\mathcal{S} = \{XX, Z\}$. A Hamiltonian is said to be stoquastic if its off-diagonal matrix entries are all nonpositive in the computational basis [Bra+08]. Bravyi and Hastings used a slightly different notion of simulation to the one we define here; as discussed in Section 4, the most important difference is that in our notion of simulation, the encoding operation must be local.

In [BH17], a sequence of 5 encodings is used to map 2-local stoquastic Hamiltonians to the transverse Ising model. We check each of the encodings in turn to see that the encodings are indeed local, so the overall result goes through with our definitions. The encodings proceed through a succession of other physical models, which we avoid defining here; see [BH17] for the details.

The encodings used are:

- TIM simulates HCD on a triangle-free graph: the encoding is the identity map.
- HCD on a triangle-free graph simulates HCB_2: the encoding attaches one additional qubit v' to each vertex v, and a qubit for each edge in the interaction graph. Each of the edge qubits is in the state $|0\rangle$, and for each vertex v, $|0\rangle_v$ is encoded as $|0\rangle_v |0\rangle_{v'}$, $|1\rangle_v$ is encoded as $|1\rangle_v |1\rangle_{v'}$. This is clearly a local encoding.
- HCB_2 simulates HCB_1: the encoding attaches $\mathrm{poly}(n)$ additional qubits, each in the state $|0\rangle$.
- HCB_1 simulates HCB_1^*: the encoding is the identity map.
- HCB_1^* simulates 2-local stoquastic Hamiltonians: the encoding maps each qubit to a subspace of two qubits in a "dual rail" encoding, and attaches some additional "mediator" qubits in a state which is a product of states of $O(1)$ qubits.

As these encodings are all local, we obtain that the transverse Ising model is a universal simulator for the class of 2-local stoquastic Hamiltonians.

To extend this simulation to k-local stoquastic Hamiltonians for $k > 2$, one can use a result from [Bra+08]. This work gave (in our terminology) a simulation of k-local termwise-stoquastic Hamiltonians with 2-local stoquastic Hamiltonians. The simulation is efficient for $k = O(1)$. A termwise-stoquastic k-local Hamiltonian H is one for which the matrices H_S occurring in the decomposition $H = \sum_S H_S$, where each subset S of subsystems on which H_S acts is of size at most k, can be taken to be stoquastic. Although all stoquastic Hamiltonians on n qubits are clearly termwise-stoquastic when viewed as n-local Hamiltonians, not all stoquastic k-local Hamiltonians are termwise-stoquastic when viewed as k-local [Bra+08]. Thus, using the simulation of [Bra+08], we obtain that the transverse Ising model is a universal simulator for stoquastic Hamiltonians, but the simulation is only efficient for termwise-stoquastic Hamiltonians.

It is shown in [CM16], using similar techniques to the proof of Theorem 41, that any family of Hamiltonians built from interactions of the form $H = \alpha Z^{\otimes 2} + A \otimes \mathbb{1} + \mathbb{1} \otimes B$, where A or B is not diagonal, can simulate TIM Hamiltonians. Thus any family of Hamiltonians of this form is also a universal stoquastic Hamiltonian simulator.

5.5 Classification of two-qubit interactions

We can complete the universality picture for two-qubit interactions by classifying the interactions into universality families. Combining the result of the previous section with Theorem 41 and a previous classification of universal classical Hamiltonians [CC16], we obtain a full classification of universality classes:

Theorem 42 *Let $\mathcal{S}$ be any fixed set of two-qubit and one-qubit interactions such that $\mathcal{S}$ contains at least one interaction which is not 1-local. Then:*

- *If there exists $U \in SU(2)$ such that U locally diagonalises $\mathcal{S}$, then $\mathcal{S}$-Hamiltonians are universal classical Hamiltonian simulators [CC16];*
- *Otherwise, if there exists $U \in SU(2)$ such that, for each 2-qubit matrix $H_i \in \mathcal{S}$, $U^{\otimes 2} H_i (U^\dagger)^{\otimes 2} = \alpha_i Z^{\otimes 2} + A_i \otimes \mathbb{1} + \mathbb{1} \otimes B_i$, where $\alpha_i \in \mathbb{R}$ and A_i, B_i are arbitrary single-qubit Hamiltonians, then $\mathcal{S}$-Hamiltonians are universal stoquastic Hamiltonian simulators [BH17; CM16];*
- *Otherwise, $\mathcal{S}$-Hamiltonians are universal quantum Hamiltonian simulators.*

We remark that the definition of universal classical simulation used in [CC16] does not quite match up with our notion of universal quantum simulation. Similarly to ours, that work associates a small number of physical qubits with each logical qubit in the simulation. However, in [CC16] the sets of physical qubits associated with distinct logical qubits are allowed to overlap. Also note that, as discussed in Section 5.4, the second (stoquastic) class of universal simulators is only efficient for termwise-stoquastic Hamiltonians.

Since the two-qubit interactions that are *not* universal must satisfy a non-trivial set of algebraic constraints, this classification immediately implies that generic two-qubit interactions are universal, an implication that can be formalised as follows:

Corollary 43 *Given any measure on the set of two-qubit Hamiltonians with full support, the subset of universal Hamiltonians has full measure.*

5.6 Spatial sparsity and simulation on a square lattice

Up to this point, we have not assumed anything about the spatial locality of the Hamiltonians we are simulating, nor the simulator Hamiltonians. Indeed, even if the target Hamiltonian has a rather simple spatial structure – for example, is a lattice Hamiltonian – this structure need not be preserved in the simulator Hamiltonian. We now show that in certain cases we can find universal simulators where all interactions take place on a square lattice. The price paid for simulating general Hamiltonians in this way (for example, those with long-range interactions) is an exponential increase in the weights required in the simulator. However, when the target Hamiltonian is spatially sparse (a class which encompasses all 2D lattice Hamiltonians), this exponential increase can be avoided.

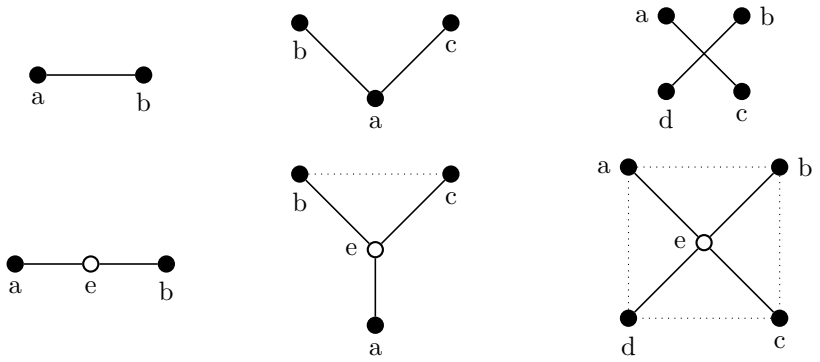
Definition 44 (Spatial sparsity [OT08]) *A spatially sparse interaction graph G on n vertices is defined as a graph in which (i). every vertex participates in $O(1)$ edges, (ii). there is a straight-line drawing in the plane such that every edge overlaps with $O(1)$ other edges and the length of every edge is $O(1)$.*

Lemma 45 *Let $\mathcal{S}$ be either $\{XX+YY+ZZ\}$, $\{XX+YY\}$ or $\{XX, Z\}$. Then any $\mathcal{S}$-Hamiltonian H on n qubits can be simulated by a $\mathcal{S}$-Hamiltonian on a square lattice of $\mathrm{poly}(n)$ qubits using weights of $O(n\Lambda_0(1/\epsilon+1/\eta))^{\mathrm{poly}(n)}$ size, where Λ_0 is the size of the largest weight in H. Furthermore if the target Hamiltonian is spatially sparse, then the weights need only be of size $O(\mathrm{poly}(n\Lambda_0(1/\epsilon+1/\eta)))$.*

Proof The final part of the statement concerning spatially sparse Hamiltonians was originally shown in [OT08] for $\{XX, Z\}$-Hamiltonians. The proof used three gadgets called fork, crossing and subdivision gadgets pictured in Figure 5, which we briefly describe here.

The subdivision gadget simulates an XX interaction between two non-interacting qubits a, b using a mediator qubit e, as pictured in Figure 5a. This can be used $O(\log k)$ times in series to simulate an interaction between two qubits separated by k qubits. The fork gadget simulates the interactions $X_aX_b + X_aX_c$ using only one interaction involving qubit a, as pictured in Figure 5b. This can be used multiple times in parallel to reduce the degree of the vertex a in the interaction graph. The crossing gadget is used to simulate $X_aX_c + X_bX_d$, for 4 qubits a, b, c, d arranged as shown in Figure 5c, via an interaction graph that has no crossings.

These gadgets can be used to simulate a spatially sparse Hamiltonian on a square lattice using only O(1) rounds of perturbation theory; we defer to [OT08] for the technical details. The gadgets were generalised for the interactions $XX+YY+ZZ$ and $XX+YY$ in [PM17], where the mediator qubit e is replaced with a pair of mediator qubits, in order to prove the result for $\{XX+YY+ZZ\}$-Hamiltonians and $\{XX+YY\}$-Hamiltonians in the same way.



(a) Subdivision gadget (b) Fork gadget (c) Crossing gadget

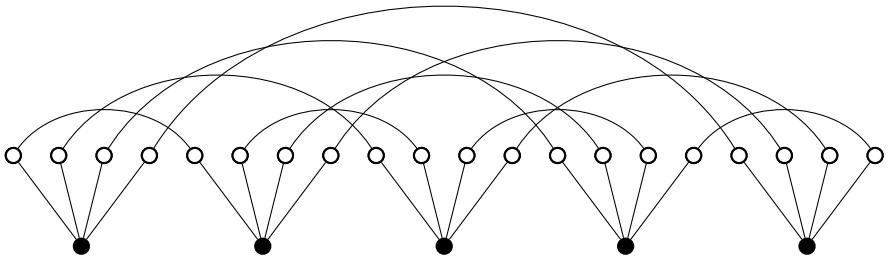
Figure 5: Subdivision, fork and crossing gadgets. In each case the top interaction pattern is simulated using the gadget underneath. White vertices denote mediator qubits with heavy 1-local terms applied.

Here we show how, if we allow more than O(1) rounds of perturbation theory, the same gadgets can be used to simulate a 2-local Hamiltonian whose interaction pattern is the complete graph on n qubits, via a simulator Hamiltonian on a square lattice of size $O(n^2) \times O(n^2)$. Any interaction graph which is a subgraph of the complete graph can easily be simulated using the same construction, simply by setting some weights to zero.

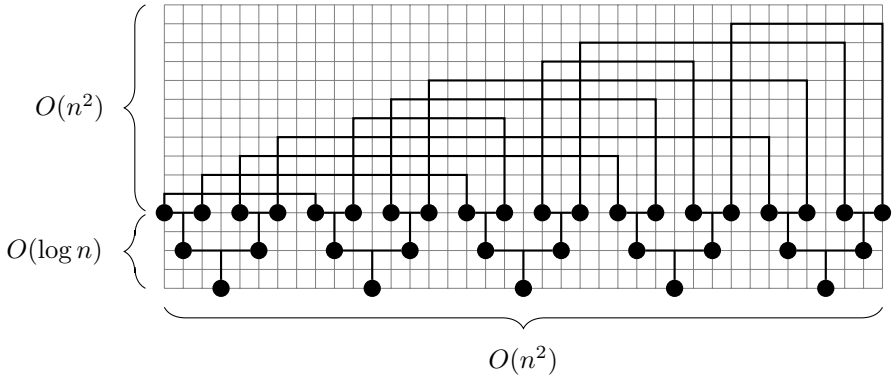
First, lay out the n qubits in a line. Each vertex in the interaction graph has $n-1$ incoming edges. Subdivide each edge just once to isolate these high degree vertices to obtain an interaction graph as shown in Figure 6a. Then using the fork gadget $O(\log n)$ times in series allows us to replace these with binary trees of depth $O(\log n)$, which can be placed directly onto a square lattice as shown in Figure 6b. The long range interactions in this graph (which are of length at most $O(n^2)$), can be fitted to the edges of the square lattice using $O(\log n)$ applications of the subdivision gadget.

At each crossing, we also need to use a crossing gadget – note that the interactions $X_a X_b$, $X_b X_c$, $X_c X_d$, and $X_d X_a$ in Figure 5c may be subdivided using a subdivision gadget so that the crossing gadget fits on the square lattice. If there is not enough space to put two crossing gadgets next to each other, then the lattice spacing can be made twice as narrow to make space. This only makes a constant factor difference to the number of qubits used and the number of rounds of perturbation theory required.

The whole procedure therefore requires a total of $O(\log n)$ rounds of perturbation theory. By Lemma 35, second-order perturbation theory requires the weights of the simulator Hamiltonian to be of size $O(\Lambda^6/\epsilon^2 + \Lambda^2/\eta^2)$, where Λ is the size of the terms H_1 and H_2. Given the simple nature of the gadgets used here, $\Lambda = O(\text{poly}(n)\Lambda_0)$ where Λ_0 is the size of the largest weight in H.



(a) First subdivide each edge to isolate each of the high degree vertices.



(b) Use the fork gadget $O(\log n)$ times at each of the high degree vertices, and lay out the resulting interaction pattern on a 2D lattice as shown above. Finally use the subdivision and crossing gadgets until the Hamiltonian is on the 2D square lattice.

Figure 6: How to simulate a Hamiltonian whose interaction pattern is the complete graph on $n = 5$ qubits with a Hamiltonian on a 2D square lattice.

Therefore r rounds of perturbation theory requires weights of size

$$\Lambda_{sim} = O\left(\text{poly}(n)\Lambda_0\left(\frac{1}{\epsilon}+\frac{1}{\eta}\right)\right)^{6^r} \tag{201}$$

Simulating the complete graph as described above requires $r = O(\log n)$, so the weights of the simulator system are $\Lambda_{sim} = (n\Lambda_0(1/\epsilon + 1/\eta))^{\text{poly}(n)}$. However, for a spatially sparse Hamiltonian simulated using only $r = O(1)$ rounds of perturbation theory as described in [OT08] the weights scale as $\Lambda_{sim} = \text{poly}(n\Lambda_0(1/\epsilon + 1/\eta))$. □

6 Consequences of universality

We finally discuss some implications of our results for quantum computation.

6.1 QMA-completeness

Oliveira and Terhal showed in [OT08] that the local Hamiltonian problem for spatially sparse qubit Hamiltonians is QMA-complete. It is observed in [CM16] that this spatially sparse Hamiltonian may be assumed to not contain any Y terms in its Pauli decomposition, by combining the work of [OT08] with a result of [BL08]. Notice that the simulations in Theorem 39 and Theorem 40 result in a spatially sparse simulator Hamiltonian if the target Hamiltonian is spatially sparse. Combined with Lemma 45, these results show that the Heisenberg interaction on a square lattice can efficiently simulate any spatially sparse qubit Hamiltonian with no Y terms, and is therefore QMA-complete. This was previously shown by Schuch and Verstraete [SV09] in the case where arbitrary 1-local terms are allowed at every site; the novelty here is that QMA-completeness still holds even if these terms are not present.

This removes the caveat of Theorem 3 in [PM17], which can now be fully stated as:

Theorem 46 *Let $\mathcal{S}$ be a set of interactions on at most 2 qubits. Assume that there does not exist $U \in SU(2)$ such that, for each 2-qubit matrix $H_i \in \mathcal{S}$, $U^{\otimes 2}H_i(U^\dagger)^{\otimes 2} = \alpha_i Z^{\otimes 2} + A_i \otimes \mathbb{1} + \mathbb{1} \otimes B_i$, where $\alpha_i \in \mathbb{R}$ and A_i, B_i are arbitrary single-qubit Hamiltonians. Then the local Hamiltonian problem for $\mathcal{S}$-Hamiltonians is QMA-complete even if the interactions are restricted to the edges of a 2D square lattice.*

Using further gadget constructions from [PM17], one can even show that the *antiferromagnetic* Heisenberg interaction is QMA-complete on a triangular lattice.

6.2 Quantum computation by simulation

We can connect universal quantum Hamiltonians to universality for quantum computation. Many constructions are now known (e.g. [JW05; Llo08; Nag12; CGW13; Tho+16; Sei+16]) which show that Hamiltonian simulation is sufficient to perform universal quantum computation. Indeed, this was already shown for

universal classical computation by Feynman [Fey85], and his construction can easily be extended to quantum computation. See [Nag08] for much more on this "Hamiltonian quantum computer" model, and many further references.

One representative example is a result of Nagaj [Nag12], who showed that for any polynomial-time quantum computation on n qubits there is a 2-local Hamiltonian H on $\mathrm{poly}(n)$ qubits with $\|H\| = O(\mathrm{poly}(n))$, a time $t = O(\mathrm{poly}(n))$, and an easily constructed product state $|\phi_0\rangle$, such that the output of the computation can be determined (with high probability) by applying e^{-iHt} to $|\phi_0\rangle$ and measuring the resulting state $|\phi_t\rangle$ in the computational basis. The description of H can be constructed in polynomial time.

Because of the strong consequences of universality, we can use any class of universal Hamiltonians to simulate an encoded version of H. Let $\mathcal{F}$ be an efficiently universal family of qubit Hamiltonians. Our definition of efficient simulation implies that, for any polynomial-time quantum computation on n qubits, there is a protocol of the following form to obtain the output of the computation:

(i). Prepare a pure state $U|\phi_0\rangle|0\rangle^{\otimes m}$ of $\mathrm{poly}(n)$ qubits, for some encoding map U such that U is a product of unitaries, each of which acts on $O(1)$ qubits.

(ii). Apply $e^{-iH't}$ for some Hamiltonian $H' \in \mathcal{F}$ such that $\|H'\| = \mathrm{poly}(n)$, and some time $t = \mathrm{poly}(n)$.

(iii). Decode the output by applying $U^\dagger$.

(iv). Measure the resulting state in the computational basis.

Observe that the first and third steps can be implemented by quantum circuits of depth $O(1)$. By universality of $\mathcal{F}$, there exists $H' \in \mathcal{F}$ such that H' is a (Δ, η, ϵ)-simulation of H for arbitrary $\epsilon > 0$. By Corollary 29, if we take $\eta, \epsilon = 1/\mathrm{poly}(n)$ and evolve according to H' for time $t = \mathrm{poly}(n)$, the resulting state $|\psi\rangle$ is distance $1/\mathrm{poly}(n)$ from an encoded version of $e^{-iHt}|\phi_0\rangle$; call that state $\mathcal{E}_{\mathrm{state}}(\phi_t)$. By Proposition 4, the expectation of any encoded measurement operator $\mathcal{E}(A)$ applied to $\mathcal{E}_{\mathrm{state}}(\phi_t)$ is the same as that of A applied to ϕ_t. Thus applying $U^\dagger$ to $\mathcal{E}(\phi_t)$ in order to undo $\mathcal{E}$, and then measuring in the computational basis, would result in the same distribution on measurement outcomes as measuring ϕ_t in the computational basis. So the distribution obtained by measuring in step (iv) is close (i.e. at total variation distance $1/\mathrm{poly}(n)$) to the distribution that would have been obtained from the measurement at the end of the simulated computation.

Thus our results show that these steps, together with time-evolution according to apparently rather simple interactions are sufficient to perform arbitrary quantum computations. For example, time-independent Heisenberg interactions with a carefully crafted pattern of coupling strengths, but no additional types of interaction, are sufficient for universal quantum computation; the same holds for XY interactions. Note that a similar statement was already known for the case of time-dependent Heisenberg interactions [DiV+00; Kem+00]: the proof

of universality there was also based on encoding, though made substantially simpler by the additional freedom afforded by time-dependence. Also note that, though not stated explicitly there, universality of the Heisenberg interaction on arbitrary graphs for quantum computation should follow from the techniques in [CGW13]. Universality of the XY interaction for quantum computation, when augmented by some additional restricted types of interactions, was shown in [CGW13; Tho+16; Sei+16].

We also showed that any universal set of 2-qubit interactions can efficiently simulate any spatially sparse Hamiltonian, even if all interactions in the simulator Hamiltonian occur on a square lattice. As there exist families of spatially sparse Hamiltonians which are universal for quantum computation (e.g. [NW08; OT08]) this implies that these interactions remain universal for quantum computation on a square lattice. For example, Heisenberg interactions are universal for quantum computation even when restricted to a 2D square lattice; as are XY interactions.

The converse perspective on this is that these Hamiltonians are more complicated to simulate than one might have previously thought. Following Lloyd's original quantum simulation algorithm [Llo96], a number of works have developed more efficient algorithms for quantum simulation, whether of general Hamiltonians [Ber+07; BCK15] or Hamiltonians specific to particular physical systems, such as those important to quantum chemistry [Has+15; Pou+15]. However, although these algorithms use very different techniques, one property which they share is that they are highly sequential; to simulate a Hamiltonian on n qubits for time t, each of the algorithms requires a quantum circuit of depth $\mathrm{poly}(n, t)$. Quantum simulation is predicted to be one of the earliest applications of quantum computers, yet maintaining coherence for long times is technically challenging. So it would be highly desirable for there to exist a Hamiltonian simulation algorithm with low depth; for example, an algorithm whose quantum part consisted of a quantum circuit of depth $\mathrm{poly}(\log(n))$.

Our results give some evidence that such a simulation algorithm is unlikely to exist, even for apparently very simple Hamiltonians such as the Heisenberg model. If there existed a Hamiltonian simulation algorithm for simulating a Heisenberg Hamiltonian on n qubits for time t, whose quantum part were depth $\mathrm{poly}(\log(n, t))$, then the quantum part of *any* polynomial-time quantum computation on n qubits could be compressed to depth $\mathrm{poly}(\log(n))$. This can be seen as a complexity-theoretic analogue of a query complexity argument [Ber+07] that lower-bounds the time to simulate an arbitrary sparse Hamiltonian. Unlilke the query complexity approach, using computational complexity theory gives evidence for hardness of simulating explicitly given local Hamiltonians. In complexity-theoretic terms, our results show that, roughly speaking[1], simulating

[1]This statement is only approximately true, for several reasons. The Hamiltonian simulation problem as we have defined it is intrinsically quantum: the task is to produce the state $e^{-iHt}|\psi\rangle$, given an input state $|\psi\rangle$. To formalise this complexity-theoretic claim, one would have to define a suitable notion of quantum reductions which encompassed such "state transformation" problems. And technically, the hardness result we prove is that the Hamiltonian simulation problem is at least as hard as PromiseBQP, the complexity class corresponding to determining whether measuring the first qubit of the output of a quantum computation is likely to return 0

any universal class of Hamiltonians is BQP-complete under QNC_0 reductions, where BQP is the complexity class corresponding to polynomial-time quantum computation, and QNC_0 is the class of depth-$O(1)$ quantum circuits.

6.3 Adiabatic quantum computation

The model of adiabatic quantum computation allows arbitrary polynomial-time quantum computations to be performed in the ground state of a family of Hamiltonians [Aha+08]. A continuously varying family of Hamiltonians $H(t)$ is used, where $0 \leq t \leq 1$. $H(0)$ and $H(1)$ are chosen such that the ground state of $H(0)$ is easily prepared, while the ground state of $H(1)$ encodes the solution to some computational problem. For example, it could be the computational history state [KSV02] encoding the entirety of a polynomial-length quantum computation. At time $t = 0$, the system starts in the ground state of $H(0)$. If the rate of change of t is slow enough, the system remains in its ground state throughout, and at time $t = 1$ the solution can be read out from the state by measuring in the computational basis. In order to perform the adiabatic computation in time $\mathrm{poly}(n)$, it is sufficient that the spectral gap of $H(t)$ is at least δ for all t, for some $\delta \geq 1/\mathrm{poly}(n)$, and that $\|H(t)\|$ and $\|\frac{d}{dt}H(t)\|$ are upper-bounded by $\mathrm{poly}(n)$ for all t [JRS07].

It was shown in [KKR06] that universal adiabatic quantum computation can be achieved using 2-local Hamiltonians. Here we argue, following a similar argument for stoquastic Hamiltonians [BH17], that any of the classes of universal Hamiltonian we have considered here can perform adiabatic quantum computation, given the ability to perform local encoding and decoding unitary operations before and after the adiabatic evolution.

Let $H(t)$ be a family of Hamiltonians used to implement an adiabatic quantum computation. For each t we define $H'(t)$ to be a (Δ, η, ϵ)-simulation of $H(t)$ using one of the previously discussed classes of universal simulators, where $\eta, \epsilon \leq n^{-c}$ for a sufficiently small constant c, and let $V(t)$ be the corresponding local isometry. From the definition of universal simulation, and the fact that the simulations increase the norm of the simulated Hamiltonian by at most a $\mathrm{poly}(n)$ factor, $H'(t)$ has spectral gap at least $\delta - 1/\mathrm{poly}(n)$ and $\|H'(t)\| = O(\mathrm{poly}(n))$. The ground state of $H'(0)$ can be prepared efficiently by applying $V(0)$ to the ground state of $H(0)$, and the ground state of $H'(1)$ can be read off efficiently by applying $V^\dagger(1)$ and measuring in the computational basis.

It remains to show that $\|\frac{d}{dt}H'(t)\| = O(\mathrm{poly}(n))$. The map $H(t) \mapsto H'(t)$ could in principle introduce singularities, as implementing an effective interaction of weight α using a second-order perturbative reduction requires weights whose scaling with α is $\alpha^{1/2}$, so for $\alpha \to 0$ the derivative becomes infinite; a similar issue applies to third-order reductions. This can be avoided, for example, by choosing a cutoff $\alpha_{\min}$, and forming a Hamiltonian $\widetilde{H}$ by replacing each weight α in the original Hamiltonian H with $\widetilde{\alpha} = \mathrm{sgn}(\alpha)\sqrt{\alpha^2 + \alpha_{\min}^2}$. If $\alpha_{\min}$ is sufficiently small (yet still inverse-polynomial in n), $\|\widetilde{H} - H\| \leq n^{-c'}$ for an arbitrarily small

or 1, given that one of these is the case. We choose to omit a discussion of these technical issues.

constant c', and also $\|\frac{d}{dt}\widetilde{H}'(t)\| = O(\text{poly}(n))$.

References

[AE11] D. Aharonov and L. Eldar. "On the complexity of commuting local Hamiltonians, and tight conditions for topological order in such systems". In: *Foundations of Computer Science (FOCS) 2011 IEEE 52nd Annual IEEE Symposium on*. IEEE. 2011, pp. 334–343.

[Aha+08] D. Aharonov et al. "Adiabatic Quantum Computation Is Equivalent to Standard Quantum Computation". In: *SIAM Review* 50.4 (2008). `quant-ph/0405098`, pp. 755–787.

[Bal05] R. C. Ball. "Fermions without fermion fields". In: *Phys. Rev. Lett.* 95.17 (2005), p. 176407.

[BCK15] D. Berry, A. Childs, and R. Kothari. "Hamiltonian simulation with nearly optimal dependence on all parameters". In: *Proc. 56th Annual Symp. Foundations Comput. Science*. 2015, pp. 792–809.

[BDL11] S. Bravyi, D. DiVincenzo, and D. Loss. "Schrieffer–Wolff transformation for quantum many-body systems". In: *Ann. Phys.* 326.10 (2011), pp. 2793–2826.

[Ber+07] D. Berry et al. "Efficient quantum algorithms for simulating sparse Hamiltonians". In: *Comm. Math. Phys.* 270.2 (2007). `quant-ph/0508139`, pp. 359–371.

[BH17] Sergey Bravyi and Matthew Hastings. "On complexity of the quantum Ising model". In: *Communications in Mathematical Physics* 349.1 (2017), pp. 1–45.

[BK02] Sergey B Bravyi and Alexei Yu Kitaev. "Fermionic quantum computation". In: *Ann. Phys* 298.1 (2002), pp. 210–226.

[BL08] J. Biamonte and P. Love. "Realizable Hamiltonians for Universal Adiabatic Quantum Computers". In: *Phys. Rev. A* 78 (2008), p. 012352.

[Bra+08] S. Bravyi et al. "The Complexity of Stoquastic Local Hamiltonian Problems". In: *Quantum Inf. Comput.* 8.5 (2008). `quant-ph/0606140`, pp. 0361–0385.

[BV05] Sergey Bravyi and Mikhail Vyalyi. "Commutative version of the local Hamiltonian problem and common eigenspace problem". In: *Quantum Inf. Comput.* 5.3 (2005), pp. 187–215.

[CC16] G. De las Cuevas and T. Cubitt. "Simple universal models capture all classical spin physics". In: *Science* 351.6278 (2016), pp. 1180–1183.

[CGW13] A. Childs, D. Gosset, and Z. Webb. "Universal Computation by Multiparticle Quantum Walk". In: *Science* 339.6121 (2013), pp. 791–794.

[CM16] T. Cubitt and A. Montanaro. "Complexity classification of local Hamiltonian problems". In: *SIAM J. Comput.* 45.2 (2016), pp. 268–316.

[Dav91] Kenneth R. Davidson. *C*-Algebras by Example*. Amer. Math. Soc., 1991.

[DiV+00] D. DiVincenzo et al. "Universal Quantum Computation with the Exchange Interaction". In: *Nature* 408 (2000). `quant-ph/0005116`, pp. 339–342.

[Fey85] R. Feynman. "Quantum mechanical computers". In: *Optics News* 11.2 (1985), pp. 11–20.

[FS14] Terence C. Farrelly and Anthony J. Short. "Causal fermions in discrete space-time". In: *Phys. Rev. A* 89.1 (2014), p. 012302.

[Has+15] M. Hastings et al. "Improving Quantum Algorithms for Quantum Chemistry". In: *Quantum Inf. Comput.* 15.1–2 (2015), pp. 1–21.

[HJ91] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1991.

[HP40] T Holstein and H Primakoff. "Field dependence of the intrinsic domain magnetization of a ferromagnet". In: *Phy. Rev.* 58.12 (1940), p. 1098.

[HTW17] V. Havlíček, M. Troyer, and J. Whitfield. *Operator Locality in Quantum Simulation of Fermionic Models*. 2017.

[HŠ03] J. Hou and P. Šemrl. "Linear Maps Preserving Invertibility or Related Spectral Properties". In: *Acta Math. Sinica* 19.3 (2003), pp. 473–484.

[JR52] N. Jacobson and C. Rickart. "Homomorphisms of Jordan Rings of Self-Adjoint Elements". In: *Trans. Amer. Math. Soc.* 72.2 (1952), pp. 310–322.

[JRS07] S. Jansen, M.-B. Ruskai, and R. Seiler. "Bounds for the adiabatic approximation with applications to quantum computation". In: *J. Math. Phys.* 48 (2007). `quant-ph/0603175`, p. 102111.

[JW05] T. Janzing and P. Wocjan. "Ergodic quantum computing". In: *Quantum Inf. Processing* 4.2 (2005). `quant-ph/0406235`, pp. 129–158.

[Kem+00] J. Kempe et al. "Theory of Decoherence-Free Fault-Tolerant Universal Quantum Computation". In: *Phys. Rev. A* 63 (2000). `quant-ph/0004064`, p. 042307.

[KKR06] J. Kempe, A. Kitaev, and O. Regev. "The complexity of the local Hamiltonian problem". In: *SIAM J. Comput.* 35.5 (2006). `quant-ph/0406180`, pp. 1070–1097.

[KSV02] A. Yu. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*. Vol. 47. Graduate Studies in Mathematics. AMS, 2002.

[Llo08] S. Lloyd. *Robustness of Adiabatic Quantum Computing*. 2008.

[Llo96] S. Lloyd. “Universal quantum simulators”. In: *Science* 273.5278 (1996), pp. 1073–1078.

[Mar67] W. S. Martindale III. “Jordan Homomorphisms of the Symmetric Elements of a Ring with Involution”. In: *J. Algebra* 5 (1967), pp. 232–249.

[Nag08] D. Nagaj. “Local Hamiltonians in Quantum Computation”. PhD thesis. MIT, 2008.

[Nag12] D. Nagaj. “Universal two-body-Hamiltonian quantum computing”. In: *Phys. Rev. A* 85 (2012), p. 032330.

[NW08] D. Nagaj and P. Wocjan. “Hamiltonian quantum cellular automata in one dimension”. In: *Phys. Rev. A* 78 (2008), p. 032311.

[OT08] R. Oliveira and B. Terhal. “The complexity of quantum spin systems on a two-dimensional square lattice”. In: *Quantum Inf. Comput.* 8 (2008), p. 0900.

[PM17] S. Piddock and A. Montanaro. “The complexity of antiferromagnetic interactions and 2D lattices”. In: *Quantum Inf. Comput.* 17.7&8 (2017), pp. 636–672.

[Pou+15] D. Poulin et al. “The Trotter Step Size Required for Accurate Quantum Simulation of Quantum Chemistry”. In: *Quantum Inf. Comput.* 15.5&6 (2015), pp. 361–384.

[Sei+16] Sahand Seifnashri et al. “Time-independent quantum circuits with local interactions”. In: *Phys. Rev. A* 93.6 (2016), p. 062342.

[SV09] N. Schuch and F. Verstraete. “Computational Complexity of interacting electrons and fundamental limitations of Density Functional Theory”. In: *Nature Phys.* 5 (2009), pp. 732–735.

[Tho+16] K. F. Thompson et al. “Time independent universal computing with spin chains: quantum plinko machine”. In: *New J. Phys.* 18.7 (2016), p. 073044.

[VC05] F. Verstraete and J. I. Cirac. “Mapping local Hamiltonians of fermions to local Hamiltonians of spins”. In: *J. Stat. Mech* 2005.09 (2005), P09012.

[WHT16] James D. Whitfield, Vojtěch Havlíček, and Matthias Troyer. *Local spin operators for fermion simulations.* 2016.