



**HAL**  
open science

## Going beyond gadgets: the importance of scalability for analogue quantum simulators

Dylan Harley, Ishaun Datta, Frederik Ravn Klausen, Andreas Bluhm, Daniel Stilck Franca, Albert H Werner, Matthias Christandl

► **To cite this version:**

Dylan Harley, Ishaun Datta, Frederik Ravn Klausen, Andreas Bluhm, Daniel Stilck Franca, et al.. Going beyond gadgets: the importance of scalability for analogue quantum simulators. Nature Communications, 2024, 15 (1), pp.6527. 10.1038/s41467-024-50744-9 . hal-04672149

**HAL Id: hal-04672149**

**<https://hal.science/hal-04672149v1>**

Submitted on 25 Aug 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

# Going beyond gadgets: The importance of scalability for analogue quantum simulators

Dylan Harley,<sup>1,\*</sup> Ishaun Datta,<sup>2</sup> Frederik Ravn Klausen,<sup>1</sup> Andreas Bluhm,<sup>3</sup>  
Daniel Stilck França,<sup>4</sup> Albert H. Werner,<sup>1</sup> Matthias Christandl<sup>1</sup>

<sup>1</sup>Department of Mathematical Sciences, University of Copenhagen,  
Universitetsparken 5, 2100 Copenhagen, Denmark

<sup>2</sup>Stanford University,  
450 Serra Mall, Stanford, CA 94305, USA

<sup>3</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG,  
38000 Grenoble, France

<sup>4</sup>Univ. Lyon, ENS Lyon, UCBL, CNRS, Inria, LIP  
Lyon Cedex 07, France

\*email: dh@math.ku.dk

---

## Abstract

Quantum hardware has the potential to efficiently solve computationally difficult problems in physics and chemistry to reap enormous practical rewards. Analogue quantum simulation accomplishes this by using the dynamics of a controlled many-body system to mimic those of another system; such a method is feasible on near-term devices. We show that previous theoretical approaches to analogue quantum simulation suffer from fundamental barriers which prohibit scalable experimental implementation. By introducing a new mathematical framework and going beyond the usual toolbox of Hamiltonian complexity theory with an additional resource of engineered dissipation, we show that these barriers can be overcome. This provides a powerful new perspective for the rigorous study of analogue quantum simulators.

## Introduction

The simulation of quantum systems has long been identified as a potential application for quantum technologies<sup>1</sup>, for which long-term benefits may range from condensed matter physics to quantum chemistry and the life sciences<sup>2,3</sup>. This problem is classically intractable, owing to exponential growth in the number of parameters required to describe the state of a many-body system, whereas the advantage of quantum hardware for this purpose is obvious: one merely has to prepare the required many-body state. On a universal quantum computer, time evolution can then be discretised and approximated by a quantum circuit, through a series of quantum gates<sup>4</sup>. This approach, known as digital quantum simulation<sup>5</sup>, has seen extensive theoretical development<sup>6,7</sup> and remains a promising route towards attaining quantum advantage<sup>8</sup>. However, useful and scalable simulations remain out of reach for near-term technology<sup>9</sup> due to the requirement of a large universal fault-tolerant quantum computer. In this work, we focus on an alternative approach: analogue quantum simulation.

Broadly speaking, an analogue quantum simulator consists of an engineered and well-controlled many-body system with adjustable interactions, with the capability to prepare initial states and perform measurements<sup>10</sup>. By tuning such a system, one aims to mimic a different target system; in this way, computing the dynamics of the target system can be accomplished through the native time evolution of the simulator, without requiring the application of a universal set of gates. These more modest

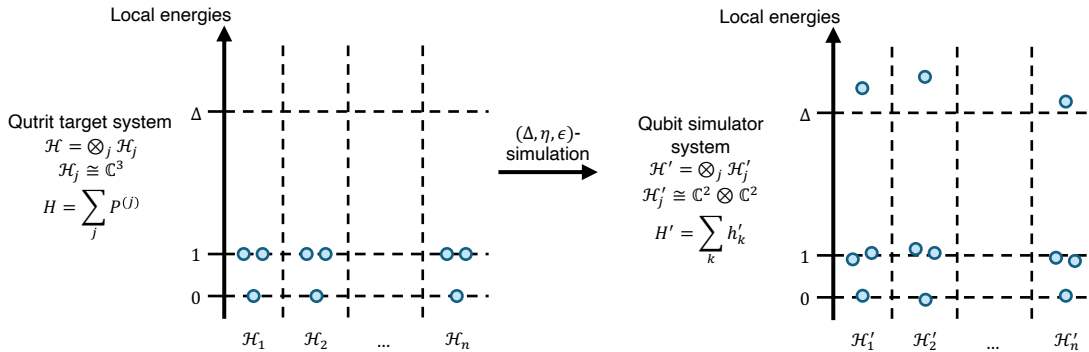


Figure 1: Qutrit-to-qubit encoding energies. A sketch of the on-site energies for a system of non-interacting qutrits under a low-energy encoding, before and after simulation in the sense introduced by Cubitt et al.<sup>11</sup>. In this example, each qutrit is mapped to a system of two qubits. The original Hamiltonian consists of a sum  $\sum_j P^{(j)}$  of rank two projectors  $P^{(j)}$  applied to each qutrit, resulting in energy levels 0,1,1. In order to simulate the qutrit Hamiltonian, an energy penalty of at least  $\Delta$  must be given to one of the four local states at each site. This is a simplification: in general the simulator sites  $\mathcal{H}'_j$  may interact and hence local energies are not well-defined.

physical requirements promise near-term potential for analogue quantum simulation, despite the inherent limitations fixed by a given experimental apparatus.

Characterisation of analogue quantum simulators is, unlike the digital case, relatively under-explored from a theoretical perspective. Existing work in this direction includes that of Cubitt et al.<sup>11</sup>, in which the authors define a notion of Hamiltonian simulation in terms of low-energy encodings: a low-energy subspace of the simulator Hamiltonian is required to approximate the spectrum of the target Hamiltonian. This notion has been extraordinarily successful in making complexity-theoretic reductions between various Hamiltonians, leading to the classification of many so-called universal families<sup>12,13,14,15,16</sup> which have the power to simulate all of many-body physics. Such reductions do not necessarily aim to capture experimental possibilities, however: as we prove, the relatively simple task of encoding a system of  $n$  non-interacting qutrits into a linear number of qubits in this regime ends up requiring a simulator system whose individual interactions scale as  $\Omega(n)$ . This scaling arises due to the dimension mismatch when one encodes a qutrit into a set of qubits, resulting in unwanted local configurations which must be prohibited in the low-energy subspace by a large energy penalty (see Fig. 1). Similar scalings are observed to arise through the use of Hamiltonian gadgets<sup>17,18</sup>, a tool used for many Hamiltonian reductions. Although the qutrit-to-qubit result does not extend to the case where the  $n$  qutrits are simulated by  $\Omega(n^2)$  qubits, we also note that blowing up the system size may in some cases necessitate strong interactions in order for correlations to spread fast enough through the enlarged system.

We argue that for practical applications to large-scale many-body simulations such as for quantum chemistry, the simulation of an  $n$ -site many-body system should not require the implementation of

individual interactions whose magnitude scales with  $n$ . The necessity of this requirement is clear from a logistical perspective, since an experimental device will only be able to implement a bounded range of energy scales on for a single interaction. However, there is also philosophical motivation to be suspicious of such scalings: a many-body Hamiltonian is inherently a modular object, and an analogue quantum simulation should reflect this. The addition of a few qubits and local interactions to one end of the physical system should require an analogous action on the simulator — it should not require the adjustment of every other interaction in the system.

Despite this additional requirement of scalability, there is also a sense in which the framework of<sup>11</sup> can be relaxed: the requirement to simulate the full physics of a target Hamiltonian in the low-energy subspace of a simulator is unnecessary in many cases. For example, one may only wish to simulate a specific set of local observables, or exploit symmetries to restrict to an invariant subspace under the Hamiltonian (a regime explored, in the case of a low-energy subspace, by Aharonov et al.<sup>19</sup>). Furthermore, as the experimental distinction between analogue and digital devices becomes increasingly blurred<sup>9,20</sup>, it is important to consider a range of experimental possibilities beyond pure Hamiltonian evolution, such as intermediate unitary pulses and open-system dynamics.

In this work, we propose a mathematical framework for analogue quantum simulators to address the above points and capture the full scope of experimentally realisable systems. We additionally develop a general characterisation for Hamiltonian gadgets, and find rigorous no-go results for their scalable use for locality reduction. Finally, we construct a new dissipative gadget which circumvents the restrictions we find in the pure Hamiltonian case. In the regime of scalable quantum simulators we do not expect to talk about a simple class of universal Hamiltonians which can simulate all others in any sense resembling previous results. On the other hand, the more general notion of simulation which we outline in this work gives rise to a new notion of universality, not phrased in terms of Hamiltonian classification but rather the dynamics of observables. We expect that here a resource theory of simulation should arise, with the power of simulators related by a partial order in analogy to the theory of multipartite states and tensor networks<sup>21,22,23</sup>.

## Results

**The analogue quantum simulator.** In this section, we describe our mathematical framework for analogue quantum simulation, for which further details can be found in the Methods section. The capabilities of a simulator are characterised by a target Hamiltonian  $H$ , a set of states  $\Omega_{\text{state}}$ , and a set of observables  $\Omega_{\text{obs}}$ ; the goal of the simulator is then to approximate the evolution of the observables in  $\Omega_{\text{obs}}$  under  $H$ , starting from initial states in  $\Omega_{\text{state}}$ . Restricting the set of states  $\Omega_{\text{state}}$  may offer practical and theoretical benefits: for example, one could reduce to those that can be reliably prepared on an experimental device, or take advantage of the symmetries of  $H$  to restrict  $\Omega_{\text{state}}$  to a specific invariant subspace and simulate only the reduced Hamiltonian. Likewise,  $\Omega_{\text{obs}}$  may reflect the capabilities of the measurement apparatus, or for a many-body system one might take advantage of a highly localised set of observables to only simulate their Lieb-Robinson light cone<sup>24</sup>, significantly reducing the hardware overhead necessary to simulate for small times. Such techniques have been studied in the context of many-body state exploration<sup>25,26</sup>, and more recently in the realm of analogue simulation<sup>27</sup>.

An analogue quantum simulator can be mathematically described in terms of three components, which we illustrate in Fig. 2. Firstly, a state encoding  $\mathcal{E}_{\text{state}}$ , which maps initial states from the target set  $\rho \in \Omega_{\text{state}}$  into the simulator system. This is defined in terms of a quantum channel, allowing one to interpret the target state as a quantum input to the simulator, in contrast to the regime of fault-tolerant digital quantum computers whose input is ultimately classical. Next, the simulator's

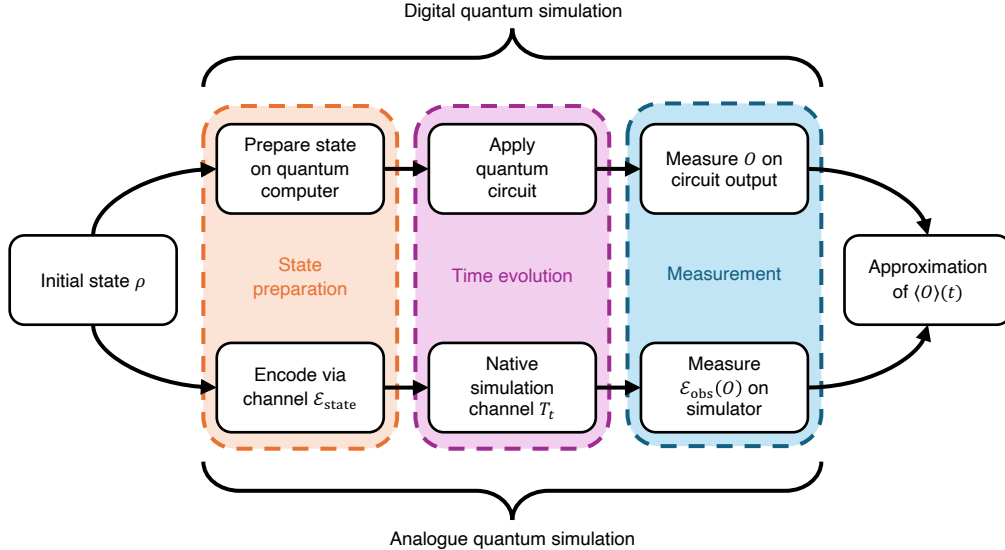


Figure 2: Analogue and digital simulation. A schematic description of our framework for analogue quantum simulation, in contrast with the digital approach. Both approaches aim to compute the observable expectation value  $\langle O \rangle(t) = \text{tr}[Oe^{-itH}\rho e^{itH}]$ , given an initial state  $\rho$ . The analogue simulator, which uses state encodings  $\mathcal{E}_{\text{state}}$  and  $\mathcal{E}_{\text{obs}}$  respectively, has accuracy  $\epsilon$  if  $|\text{tr}[Oe^{-itH}\rho e^{itH}] - \text{tr}[\mathcal{E}_{\text{obs}}(O)(T_t \circ \mathcal{E}_{\text{state}})(\rho)]| \leq \epsilon$ .

time evolution is specified by a family of quantum channels  $\{T_t\}_{t \in [0, t_{\text{max}}]}$ . These describe the dynamics of the simulator, for example the evolution under a simulator Hamiltonian  $H'$ . However, one could also consider  $T_t$  accounting for interactions with a bath (such capability is required by the criteria for analogue simulators given by Cirac et al.<sup>10</sup>), modelling errors, or capturing other engineered controls reflecting the possibilities of the experimental apparatus. The final component of the simulator is an encoding for observables  $\mathcal{E}_{\text{obs}}$ , a unital and completely positive map which sends an observable of interest  $O \in \Omega_{\text{obs}}$  to the relevant observable to be measured on the simulator system.

After the encoding and time evolution steps for a given initial state  $\rho \in \Omega_{\text{state}}$  and time  $t \in [0, t_{\text{max}}]$ , the simulator system lies in the state  $(T_t \circ \mathcal{E}_{\text{state}})(\rho)$ , upon which one measures the encoded observable  $\mathcal{E}_{\text{obs}}(O)$  for a chosen  $O \in \Omega_{\text{obs}}$ . The simulation has accuracy  $\epsilon$  if the expectation value of this measurement is within  $\epsilon$  of its target value — that is, the expected value of measuring  $O$  on  $e^{-itH}\rho e^{itH}$ . Note that rather than a Hamiltonian  $H$ , one could just as easily consider the simulation of an open target system, for instance described by a quantum dynamical semigroup<sup>28,29</sup>.

For practical simulators, some constraints must be placed on the maps used in this definition. We generally assume that both  $\mathcal{E}_{\text{state}}$  and  $\mathcal{E}_{\text{obs}}$  are local in a sense we define, to ensure that local errors correspond to local noise on the target system, and that local observables can be measured locally on the simulator system. Moreover, the time evolution channel  $T_t$  should be implementable without the need for feed-forward measurement results for adaptive control: the lack of error correction is an important and characteristic feature of analogue simulators.

**Generalising Hamiltonian gadgets.** A ubiquitous technique for Hamiltonian reductions in complexity theory is the use of so-called Hamiltonian gadgets<sup>30,17,31,18</sup>. These provide a recipe to simulate complicated many-body interactions from a more restrictive family, for example to simulate a 3-body

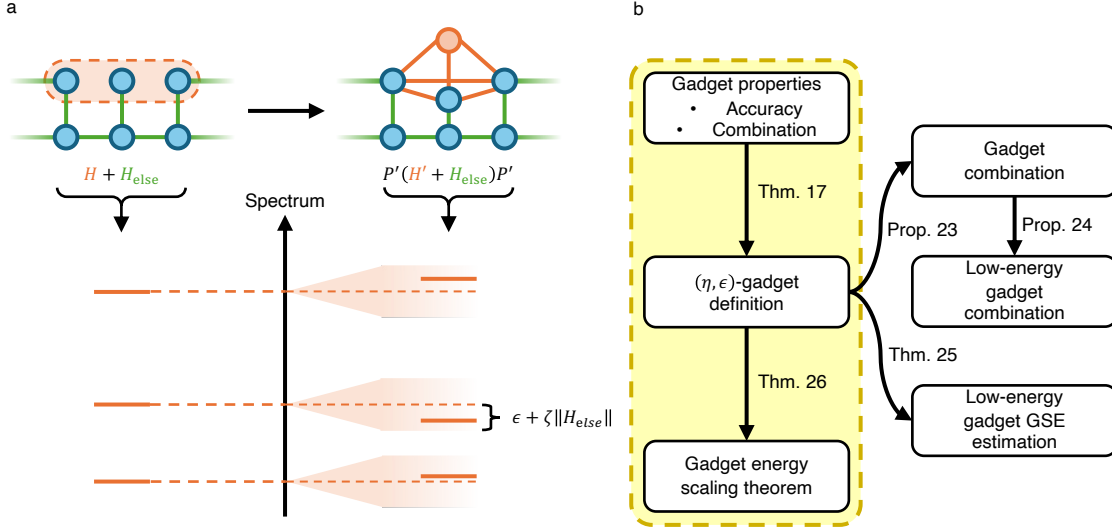


Figure 3: Hamiltonian gadget characterisation. (a) The interaction hypergraph of a Hamiltonian containing a 3-local interaction which is replaced by a 2-local gadget. The gadget property requires that the spectrum is unchanged up to  $\epsilon + \zeta \|H_{\text{else}}\|$ , for  $\epsilon, \zeta > 0$ , when restricted by a projector  $P'$ . (b) Structure of gadget results; boxes highlighted in yellow indicate the central argument for the energy scaling no-go. We first formalise the desirable properties of gadgets and show that they imply a general definition, from which we can prove the energy scaling theorem along with various combination properties, including a generalisation of a result of Bravyi et al.<sup>31</sup> for ground state energy (GSE) estimation.

interaction using 2-body interactions. In this section we arrive at a general formalism for such constructions, in order to prove that they are associated with unavoidable energy scalings which pose a significant challenge for experimental realisations. The usual procedure, as formulated by Bravyi et al.<sup>18</sup> for example, is as follows: starting from a target Hamiltonian  $H$  (which might be a single interaction in a far larger system), one first adjoins an ancillary qubit  $m$ . On this enlarged system one defines the gadget Hamiltonian by  $H' = \Delta|1\rangle\langle 1|_m + V$ , where  $\Delta \gg 1$  is a large parameter used to define a low-energy subspace approximately in terms of the  $|0\rangle$  state of the mediator qubit, and  $V$  is a relatively small term which, via a perturbative approximation, effectively simulates the target  $H$  in this low-energy subspace.

It is not surprising that this method of construction generically requires strong interactions corresponding to the large value of  $\Delta$  needed to provide a sufficiently high energy penalty, but it is not immediately clear that there is no way around this cost (possibly outside of the perturbative regime). Indeed, several works<sup>32,33,34</sup> have explored the optimisation of Hamiltonian gadgets for practical implementation, though generally the problem scaling interactions is not eliminated entirely. In this work we produce a generalised framework for gadgets in order to prove a lower bound for such scalings, suggesting that such techniques may be unsuitable for experiments on large systems. Our results are summarised in Fig. 3(b), and full mathematical details can be found in the Methods section.

Let  $H$  be the target Hamiltonian on a Hilbert space  $\mathcal{H}$ , and let the gadget Hamiltonian  $H'$  act on the space  $\mathcal{H} \otimes \mathcal{A}$ , for  $\mathcal{A}$  some ancillary system. We require the following two properties of  $H'$ , illustrated in Fig. 3(a):

- Accuracy: The spectrum of  $H$  should be approximated by that of  $H'$ , in some subspace defined by a projector  $P' \in \text{Proj}(\mathcal{H} \otimes \mathcal{A})$ , up to error  $\epsilon \geq 0$ .
- Combination: The above property should hold even when any additional Hamiltonian  $H_{\text{else}}$  is added to both  $H$  and  $H'$ , at the expense of an additional spectral error  $\zeta \|H_{\text{else}}\|$ , where  $\|\cdot\|$  denotes the operator norm.

For small error parameters  $\epsilon$  and  $\zeta$ , these properties are — non-trivially — sufficient to force  $H'$  to satisfy the following definition, which resembles previously studied notions of simulation<sup>11,18</sup>. We say that  $H'$  is an  $(\eta, \epsilon)$ -gadget (where  $\eta$  is a new parameter related to  $\zeta$ , also measuring the ability of the gadget to combine with other terms) for  $H$  if there exists a projector  $P \in \text{Proj}(\mathcal{A}) \setminus \{0\}$  and a unitary operator  $U \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  satisfying two conditions. Firstly,  $U$  must be  $\eta$ -close to the identity:  $\|U - \text{I}\| \leq \eta$ . Then, defining the projector  $P'$  by  $P' = U(\text{I} \otimes P)U^\dagger$  (so that  $\eta$  in some sense quantifies how close  $P'$  is to a pure projection on the ancillary system), the second condition ensures that the spectrum  $P'H'P'$  should approximate that of  $H$ , up to some multiplicity:  $\|P'H'P' - U(H \otimes P)U^\dagger\| \leq \epsilon$ .

This gadget definition expresses the quality of a gadget through two parameters:  $\epsilon$  can be thought of as the absolute error of the gadget, whilst  $\eta$  bounds the error incurred when the gadget is combined with other interactions in a Hamiltonian. In particular, when  $\|H_{\text{else}}\| \sim n$  grows with the size of the system,  $\eta$  must correspondingly shrink to hold the error constant.

Despite the generality of this definition, it is sufficient to guarantee that such gadgets can be combined in parallel. That is, given a many-body Hamiltonian  $H = \sum_i H_i$  and sufficiently good gadgets  $H'_i$  for each of the individual terms  $H_i$ , the Hamiltonian  $H' = \sum_i H'_i$  constitutes a good gadget for  $H$ . A similar result holds for low-energy gadgets (for which the projector  $P'$  is replaced by a projector onto the low-energy subspace of  $H'$ ), and also leads to a generalisation of the ground state energy estimation result of Bravyi et al.<sup>31</sup>.

On the other hand we show that, when used for certain types of reduction, gadgets come at an unavoidable energy cost. In particular, any attempt to simulate a  $k$ -body interaction  $H$  via a gadget  $H'$  consisting of  $k'$ -body interactions for  $k' < k$  necessarily requires interaction strengths scaling as  $\Omega(\eta^{-1})$ . In order to control the absolute error of a many-body system,  $\eta^{-1}$  must grow with the size of the system, leading to unfeasible energy scalings and constituting a significant barrier for Hamiltonian reductions in the regime of experimentally realisable analogue quantum simulators.

**Gadgets from the quantum Zeno effect.** To circumvent our no-go result for scalable Hamiltonian locality reduction, in this section we exhibit a new kind of gadget, taking advantage of the non-unitary possibilities afforded by a general simulation channel  $T_t$ . This works by restricting the mediator qubit to its  $|0\rangle$  state not with a strong interaction but through inertia induced by the quantum Zeno effect. This powerful resource can be used to build a direct  $k$ -to- $(\lceil k/3 \rceil + 1)$ -local gadget, without interactions scaling with the size of the system.

The recipe for this construction is qualitatively similar to that for usual Hamiltonian gadgets: starting from a target interaction  $H$ , a mediator qubit  $m$  is adjoined to the system, and evolves under a simulator Hamiltonian  $H'$ . In this case, however,  $H'$  need not contain an interaction  $\Delta|1\rangle\langle 1|_m$ , with  $\Delta$  scaling with the size of the system. Instead, a dissipative channel is applied to the qubit  $m$  at regular intervals separated by time  $\delta t$ . Provided that  $\delta t$  is small enough, the quantum Zeno effect keeps  $m$  effectively fixed in its  $|0\rangle$  state with high probability, whilst the remainder of the system evolves as though under the target Hamiltonian  $H$ . This is illustrated by Fig. 4, and a rigorous description can be found in the Methods.

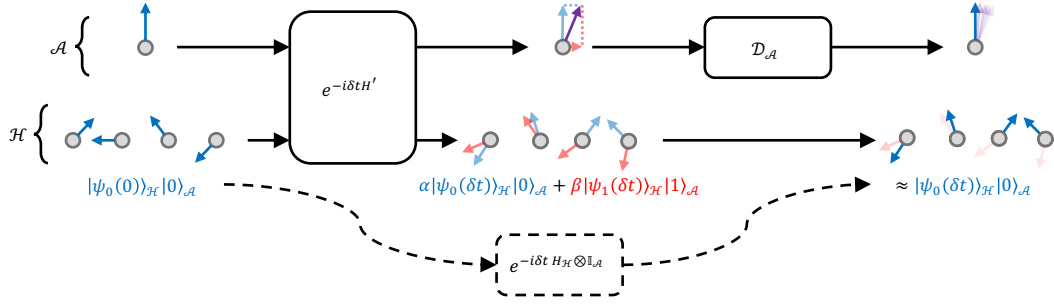


Figure 4: Dissipative gadget evolution. Circuit representation of the non-unitary gadget procedure for a single timestep. The initial state  $|\psi_0(0)\rangle_{\mathcal{H}}|0\rangle_{\mathcal{A}}$  is evolved under a Hamiltonian  $H'$  for time  $\delta t$ , resulting in a superposition of states with the ancillary qubit in the  $|0\rangle_{\mathcal{A}}$  and  $|1\rangle_{\mathcal{A}}$  positions. After applying the dissipative channel to the  $\mathcal{A}$  system, the system collapses to its  $|0\rangle_{\mathcal{A}}$  state with high probability due to the quantum Zeno effect. Meanwhile, the resulting state on the  $\mathcal{H}$  system approximately corresponds to evolution under a different Hamiltonian,  $H_{\mathcal{H}}$ .

These non-unitary gadgets may be combined with other terms in a Hamiltonian at no extra cost (effectively corresponding to a gadget with combination error parameter  $\eta = 0$ ), yielding an improvement on any possible pure Hamiltonian gadget. On the other hand, the construction has various caveats: strong interactions (though not scaling with system size) are still necessary for high accuracy of a single gadget, and we expect that combining multiple such gadgets will require strong interactions, to suppress the probability of any ancillary qubit transitioning into the  $|1\rangle$  state. Moreover, the precisely engineered stroboscopic dissipation channel constitutes a new experimental challenge.

Nevertheless, this construction provides insight into the ways in which non-unitary dynamics might be exploited for practical analogue quantum simulation problems — indeed, similar tools have already found applications in theory<sup>35,36</sup> and in practice<sup>37</sup> for digital quantum computing. In light of our theorems implying extensive interaction scaling for qutrit-to-qubit mappings and gadget locality reduction, which effectively serve as no-go theorems for practical universal simulators built from pure Hamiltonian dynamics, we anticipate that similar hybrid techniques will constitute a powerful tool for attaining useful quantum advantage with quantum simulators.

## Methods

### Criteria for quantum computation and simulation

In a review of the prospective possibilities of quantum computing<sup>38</sup> the author provided a set of requirements, now known as the DiVincenzo criteria, designed to serve as a full specification for implementations of universal quantum computers. These are summarised in Fig. 5.

As well as concretely providing the experimentalist with a necessary set of criteria to aim towards, the sufficiency of the DiVincenzo criteria provides the theorist with a canonical yardstick to judge the applicability of their protocol to idealised quantum hardware. It is therefore important that such requirements reflect exactly what can be expected from quantum technology in the long term, neither excluding feasible technologies nor including unfeasible procedures.



The DiVincenzo criteria for quantum computation	The Cirac-Zoller criteria for quantum simulation
A. A scalable physical system with well-characterised qubits.	I. A system of bosons/fermions confined in a region in space, containing a large number of degrees of freedom.
B. The ability to initialise the state of the qubits to a simple fiducial state, such as $ 000 \dots\rangle$ .	II. The ability to approximately prepare a known quantum state – ideally pure.
C. Long relevant decoherence times, much longer than the gate operation time.	III. The ability to engineer and adjust the values of a set of interactions between the particles, and possibly external fields or a reservoir.
D. A universal set of quantum gates.	IV. The capability to measure the system, either on specific sites or collectively.
E. A qubit-specific measurement capability.	V. A procedure for increasing confidence in the results of the simulation.

Figure 5: A summary of the DiVincenzo<sup>38</sup> and Cirac-Zoller<sup>10</sup> criteria. The DiVincenzo criteria provide necessary and sufficient requirements for universal digital quantum computers. Similarly, the Cirac-Zoller criteria offer a set of requirements for analogue quantum simulation, for which universality may not be available.

A similar set of criteria for analogue quantum simulators is discussed by Cirac et al.<sup>10</sup>, also summarised in Fig. 5. These are all natural requirements to ask of a quantum simulator, but it is noteworthy that criterion III does not provide any restriction on the interactions that one should expect the simulator to include. This leads to a problem which does not arise for the DiVincenzo criteria: whereas a quantum computer can approximate arbitrary  $k$ -qubit gates from the compact set  $U((\mathbb{C}^2)^{\otimes k})$  of unitary transformations relatively cheaply due to the Solovay-Kitaev theorem<sup>39</sup>, the task of an analogue quantum simulator is to implement  $k$ -qudit interactions from the unbounded set of possible Hamiltonians  $\text{Herm}((\mathbb{C}^d)^{\otimes k})$ . The ability to realise arbitrarily strong interactions on a physical device is clearly an impossibility.

Thus, the key extra criterion which we demand of an analogue quantum simulator is that the encoding of the target Hamiltonian should be size-independent. Concretely, if the Hamiltonian  $H$  to be encoded consists of local interactions  $(h_i)_{i=1}^m$  on  $n$  sites then the encoding of individual terms should not depend, for instance by polynomial scaling of interaction strengths, on the size of the physical system  $n$ . In particular, we argue that methods for practical analogue quantum simulation must respect a limit on the interaction strengths of the simulator Hamiltonian. The strongest interactions should be bounded by some constant fixed by physical limitations, and the weakest interactions should be similarly bounded from below (since sufficiently weak interactions will be overwhelmed by noise in the simulator). In addition, in order to ensure the local and size-independent encoding of each site into the simulator, we argue that the the simulator should grow no faster than linearly with the size of the target system. If each site is encoded into more than  $O(1)$  simulator sites, it will be impossible to encode the full system into a simulator of the same dimension while preserving geometric locality (without introducing scaling interactions). We summarise these requirements with the following qualitative definition:

**Definition 1** (Size-independent simulation). *We say that an analogue quantum simulation is size-independent if the simulation of a  $n$ -site Hamiltonian can be implemented scalably with  $n$ . By this, we mean that the number of qubits used in the simulation should grow no faster than linearly in  $n$ , and the interaction strengths necessary should remain  $\Theta(1)$ .*

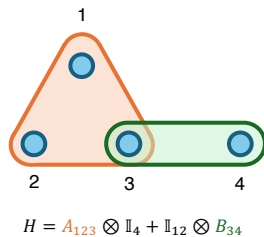


Figure 6: Example Hamiltonian interaction hypergraph. A Hamiltonian  $H$  on 4 qubits, and its associated interaction hypergraph. The Hamiltonian consists of a 3-local (orange) and a 2-local (red) term, so we say that  $H$  is 3-local.

It is worth noting that further formalisation is required to make this definition robust. For example, suppose we are given a Hamiltonian  $H = h_1 + h_2$  where  $\|h_1\|, \|h_2\| = O(n^{-1})$ , which violates the size-independence requirement. One could simply define  $h'_1 = h_1 + K$ ,  $h'_2 = h_2 - K$ , for some  $K = \Theta(1)$ , and then  $H = h'_1 + h'_2$  can be written in a form which does not obviously violate Definition 1. To exclude such possibilities, we could impose an additional requirement that  $H$  is given in a canonical form, such as that described by Wilming et al.<sup>40</sup>.

As well as being experimentally and qualitatively desirable, encoding interactions independently has quantitative benefits; as noted by Cubitt et al.<sup>11</sup>, for a suitably local Hamiltonian encoding, local errors on the simulator system will correspond to local errors on the target system. For NISQ hardware, this represents an extremely useful way to mitigate the negative effects of a noisy simulation: rather than random scrambling, noise can be viewed as the manifestation of physically reasonable noisy effects on the target system.

Finally, studying the power of Hamiltonians subject to interaction energies that are constant in system size is well-motivated in its own right, from the perspective of Hamiltonian complexity. For example, Aharonov et al.<sup>19</sup> show that restriction to such Hamiltonians will necessarily sacrifice some sense of universality of the simulator. Earlier results in Hamiltonian complexity theory<sup>31</sup>, however, show that in many cases it is still possible to simulate ground state energies up to an extensive error.

## Hamiltonian complexity theory

We say that a Hamiltonian  $H$  on the space of  $n$  qubits  $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$  is  $k$ -local if it can be written as  $H = \sum_{j=1}^N h_j$ , where each of the terms  $h_j$  acts on at most  $k$  of the qubit sites. We consider the  $h_j$  individual interactions in the Hamiltonian and make reference to the interaction hypergraph, whose vertices are qubits and whose (hyper)edges are interactions (joining the qubits on which they act), illustrated in Fig. 6.

Informally, the  $k$ -local Hamiltonian problem asks whether the ground state energy of a  $k$ -local Hamiltonian is less than  $a$ , or greater than  $b$ , for some real numbers  $a < b$  separated by a suitably large gap. This problem lies in the QMA complexity class: the natural quantum analogue to the classical NP, containing problems whose solutions can be efficiently verified (but not necessarily found) on a quantum computer.

**Definition 2** (*k*-local Hamiltonian problem). *The k-local Hamiltonian problem is the promise problem which takes as its input a k-local Hamiltonian  $H = \sum_{j=1}^N h_j$  on the space of  $n$  qubits  $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ , where  $N = \text{poly}(n)$ , and for each  $j$  we have  $\|h_j\| \leq \text{poly}(n)$  and  $h_j$  is specified by  $O(\text{poly}(n))$  bits.*

*Given  $a < b$  with  $b - a > 1/\text{poly}(n)$ , let  $\lambda_0(H)$  denote the lowest eigenvalue of  $H$ . Then the output should distinguish between the cases*

- *Output 0: The ground state energy of  $H$  has  $\lambda_0(H) \leq a$ .*
- *Output 1: The ground state energy of  $H$  has  $\lambda_0(H) \geq b$ .*

Through the Feynman-Kitaev circuit-to-Hamiltonian construction<sup>41</sup>, it was established that the 5-local Hamiltonian problem is QMA-complete, and subsequent works optimising the construction<sup>30</sup> and using gadget techniques<sup>42</sup> reduced this further to show the QMA-completeness of the 2-local Hamiltonian problem. Various further optimisations have been found to refine the problem and further restrict the family of allowed Hamiltonians (see for example<sup>43,44</sup>); indeed hardness results have been shown to hold even under the significant restriction to 1-dimensional translationally invariant systems<sup>45</sup>. QMA-completeness is closely related to a notion of universality for simulators; an equivalence was proved by Kohler et al.<sup>15</sup>.

The constructions involved in the aforementioned results contain Hamiltonian interaction strengths which scale polynomially, or exponentially, with system size — such Hamiltonians are infeasible for an analogue simulator. A notable exception to this is the work of Bravyi et al.<sup>31</sup>, in which the authors use the Schrieffer-Wolff transformation to show that bounded-strength interactions are sufficient for one to reproduce the ground-state energy of the original Hamiltonian up to an extensive error.

As much of this Hamiltonian simulation literature focuses on specific complexity-theoretic problems, comparatively little work has been done to actually define a mathematical framework for analogue quantum simulation to be used in experiment. Notable recent work in this direction includes that of Cubitt et al.<sup>11</sup>, in which the authors study methods of encoding Hamiltonians via a map  $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \rightarrow \text{Herm}(\mathcal{H}')$ , which satisfy the natural requirement of preserving the spectrum of observables. Additionally, in the case that  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$  is a space of many sites, they introduce the further notion of local encodings, which map local observables in  $\mathcal{H}$  to local observables in  $\mathcal{H}' = \otimes_{i=1}^{n'} \mathcal{H}'_i$ . By deriving the most general possible form of a spectrum-preserving Hamiltonian encoding, and then imposing natural locality conditions, the authors arrive at the following definition.

**Definition 3** (Local Hamiltonian encoding<sup>11</sup>). *A local Hamiltonian encoding is a map  $\mathcal{E}_{\text{obs}} : \text{Lin}(\otimes_{i=1}^n \mathcal{H}_i) \rightarrow \text{Lin}(\otimes_{i=1}^{n'} \mathcal{H}'_i)$  of the form*

$$\mathcal{E}_{\text{obs}}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger, \quad (1)$$

*where  $P$  and  $Q$  are locally distinguishable orthogonal projectors on an ancillary space  $\mathcal{A} = \otimes_{i=1}^n \mathcal{A}_i$ , and  $V = \otimes_{i=1}^n V_i$  where  $V_i \in \text{Isom}(\mathcal{H}_i \otimes \mathcal{A}_i, \mathcal{H}'_i)$  for all  $i$ . Here  $\bar{M}$  denotes the complex conjugate of the matrix  $M$ .*

Projectors  $P, Q \in \text{Proj}(\otimes_i \mathcal{A}_i)$  are locally distinguishable if, for all  $i$ , there exist orthogonal projectors  $P_i, Q_i \in \text{Proj}(\mathcal{A}_i)$  such that  $(P_i \otimes \mathbb{I})P = P$  and  $(Q_i \otimes \mathbb{I})Q = Q$ . Generally, we consider the case of  $\text{rank}(P) > 0$  (referred to as standard<sup>11</sup>), for which one can define a corresponding state encoding

$$\mathcal{E}_{\text{state}}(\rho) = V(\rho \otimes \tau)V^\dagger, \quad (2)$$

where  $\tau$  is a state on  $\mathcal{A}$  satisfying  $P\tau = \tau$ .

Moreover, the authors define the following notion of simulation, which relaxes the requirements of locality and allows for some error in the simulated eigenvalues.

**Definition 4** ( $(\Delta, \eta, \epsilon)$ -simulation<sup>11</sup>). *A Hamiltonian  $H' \in \text{Herm}(\mathcal{H}') = \text{Herm}(\otimes_{i=1}^n \mathcal{H}'_i)$  is said to  $(\Delta, \eta, \epsilon)$ -simulate a Hamiltonian  $H \in \text{Herm}(\mathcal{H}) = \text{Herm}(\otimes_{i=1}^n \mathcal{H}_i)$  if there exists a local encoding (Definition 3)  $\mathcal{E}_{\text{obs}}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$  such that*

(i) *There exists an encoding  $\tilde{\mathcal{E}}_{\text{obs}}(M) = \tilde{V}(M \otimes P + \bar{M} \otimes Q)\tilde{V}^\dagger$  (where  $\tilde{V} \in \text{Isom}(\mathcal{H} \otimes \mathcal{A}, \mathcal{H}')$  need not have a tensor product structure as in Definition 3) such that  $\|V - \tilde{V}\| \leq \eta$  and  $\tilde{\mathcal{E}}_{\text{obs}}(\mathbb{I}) = P_{\leq \Delta(H')}$  is the projection onto the low-energy ( $\leq \Delta$ ) subspace of  $H'$ , and*

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

This approach (later generalised by Apel et al.<sup>46</sup> and refined with resource constraints by Zhou et al.<sup>16</sup>) provides an elegant framework to capture a notion of one Hamiltonian fully simulating another. However, we believe that this regime does not capture the scope of possibilities for analogue quantum simulation experiments. On one hand, the formalism requires the entire physics of the target system to be encoded into the low-energy subspace of a simulator — this rules out simulators which only simulate part of the target system, or in a different subspace. On the other hand, the formalism is too broad in the sense that it does not prohibit unrealistically scaling interaction strengths in violation of Definition 1.

## Framework

The generic task of an analogue quantum simulator is to estimate the dynamics of observables in a system  $\mathcal{H}$  under the evolution of a target Hamiltonian  $H$ , up to some maximum time  $t_{\text{max}}$ . In particular, it is not always necessary to simulate the entire target system in arbitrary configurations: it may be convenient to restrict to a particular subset of initial states  $\Omega_{\text{state}}$ , for example lying in a subspace invariant under the Hamiltonian or corresponding to the states which can be reliably prepared by the simulator, and similarly to a particular subset of observables of interest  $\Omega_{\text{obs}}$ . We denote by  $\mathcal{H}'$  the Hilbert space corresponding to the simulator system, and for  $t \in [0, t_{\text{max}}]$  we write  $T_t : D(\mathcal{H}') \rightarrow D(\mathcal{H}')$  for the family of time evolution quantum channels implemented by the simulator, where  $D(\mathcal{H}')$  is the set of density matrices on  $\mathcal{H}'$ . This approach, in which we view simulations in terms of individual observables rather than the entire Hamiltonian, has been considered in earlier work<sup>25,26,27</sup>.

The minimal requirement for a simulator is that it should approximate the expectation values of the elements of  $\Omega_{\text{obs}}$ . That is,  $\text{tr}[Oe^{-iHt}\rho e^{iHt}]$  should be close to  $\text{tr}[O'T_t(\rho')]$  for all  $\rho \in \Omega_{\text{state}}$  and  $O \in \Omega_{\text{obs}}$ , where  $\rho'$  and  $O'$  are some encoded versions of the states and operators respectively. Notice that, in principle, the experimentalist could be using a completely different simulator for each choice of  $\rho$  and  $O$ , with  $\mathcal{H}'$  a space large enough to contain all of them and by encoding  $\rho$  into several copies. However, this would violate the size-independence requirement of Definition 1 if  $\Omega_{\text{obs}}$  and  $\Omega_{\text{state}}$  both do not only contain  $O(1)$  elements. Furthermore, it is natural to consider analogue quantum simulators as machines taking quantum, rather than classical, input — possibly prepared by another experiment — which cannot be cloned. For this reason, we assume that the state encoding takes the form of a quantum channel  $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \rightarrow D(\mathcal{H}')$ . Correspondingly, to accommodate for quantum outputs, we require the observable encoding  $O \mapsto O'$  to be a unital and completely positive map  $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \rightarrow \text{Herm}(\mathcal{H}')$ . This ensures that the Hilbert-Schmidt dual operator  $\mathcal{E}_{\text{obs}}^*$  is a quantum channel, so measurement of

$\mathcal{E}_{\text{obs}}(O)$  on  $\rho'$  can equivalently be thought of as a measurement of  $O$  on a decoded state  $\mathcal{E}_{\text{obs}}^*(\rho)$ . This perspective sets analogue quantum simulators apart from the framework of digital quantum computation, for which fault-tolerant architectures require both inputs and outputs to be classical.

This definition is still sufficiently versatile to capture the simulation of global observables that are a sum of local parts  $O = \sum_k O_k$  (a task, for example, useful for variational quantum algorithms<sup>47</sup>), in the following way. Often the  $O_k$  cannot be simultaneously measured due to non-commutativity relations or experimental limitations. The simplest approach to estimating  $O$  is to run many simulations, measuring one of the  $O_k$  each time (this process can be sped up by combining simultaneously measurable terms<sup>48</sup>), and summing the average results.

The above discussion leads us to the following definition, which is illustrated by Fig. 2.

**Definition 5** (Analogue quantum simulation). *Given a set of states  $\Omega_{\text{state}}$  on a Hilbert space  $\mathcal{H}$ , a normalised set of observables  $\Omega_{\text{obs}}$  (i.e.  $\|O\| = 1$  for all  $O \in \Omega_{\text{obs}}$ , where  $\|\cdot\|$  denotes the operator norm), a time  $t_{\text{max}} > 0$ , a Hamiltonian  $H \in \text{Herm}(\mathcal{H})$ , and  $\epsilon > 0$ , we say that a family of quantum channels  $T_t : D(\mathcal{H}') \rightarrow D(\mathcal{H}')$ , for  $t \in [0, t_{\text{max}}]$  simulates  $H$  with respect to  $\Omega_{\text{state}}$  and  $\Omega_{\text{obs}}$  with accuracy  $\epsilon$  if there exists*

1. A state encoding quantum channel  $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \rightarrow D(\mathcal{H}')$  which maps states to the simulator Hilbert space  $\mathcal{H}'$ ,
2. An observable encoding, given by a unital and completely positive map  $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \rightarrow \text{Herm}(\mathcal{H}')$ ,

such that

$$\left| \text{tr}[\mathcal{E}_{\text{obs}}(O)(T_t \circ \mathcal{E}_{\text{state}})(\rho)] - \text{tr}[O(e^{-itH}\rho e^{itH})] \right| \leq \epsilon, \quad (3)$$

for all  $\rho \in \Omega_{\text{state}}$ ,  $O \in \Omega_{\text{obs}}$ , and  $t \in [0, t_{\text{max}}]$ .

Our use of a Hamiltonian  $H$  for the target system is mostly for simplicity; the simulation of more general dynamics, of open quantum systems for example, can be defined analogously, with the target Hamiltonian  $H$  replaced by any generator of a quantum dynamical semigroup<sup>28,29</sup>. It should be noted also that Definition 5 could equivalently have been phrased in terms of a set of POVMs rather than observables  $\Omega_{\text{obs}}$ . We use the latter for convenience in relating our work to other results. It is plausible that one could engineer a time-dependent observable encoding  $\mathcal{E}_{\text{obs}}$ , but here we restrict our focus to the time-independent case to avoid the complexity of the simulation task being hidden in this step.

By the triangle inequality, (3) holds for any convex combination of the states and observables in  $\Omega_{\text{state}}$  and  $\Omega_{\text{obs}}$  respectively, so we could without loss of generality assume that the two sets are convex to begin with.

Often the simulation channels  $T_t$  in Definition 5 are taken simply as time evolution under some simulator Hamiltonian  $H' \in \text{Herm}(\mathcal{H}')$ , but it is useful to consider a more general case. Firstly, this allows one to directly account for, and possibly exploit, dissipative errors in the experimental setup<sup>49</sup>. Secondly, it enables the possibility of a more complicated simulation experiment, for example involving intermediate measurements. Moreover, it is important to allow the simulation of open quantum systems for our definition to be consistent with criterion III of Fig. 5. Despite the generality afforded by Definition 5, we emphasise that experimentally practical simulations should be size-independent as in Definition 1. That is, the implementation of  $T_t$  should not require engineering a system of size which grows more than

linearly in  $n$ , or boundlessly scaling interaction energies. Another important constraint is that  $T_t$  should not include the use of adaptive channels based on feed-forward measurements — hence distinguishing the process from digital quantum computation.

We note that Hamiltonian models of quantum computation such as quantum walks<sup>50</sup> and previous notions of dynamical Hamiltonian simulation<sup>51</sup> are not consistent with our definition of analogue simulation: such constructions also incur scalings in both the system size and in necessary evolution time (corresponding to scalings in interaction strength, if time is normalised) which violate the size-independence conditions of Definition 1.

## Local encodings

Although Definition 5 is phrased in terms of general encoding maps, it is practically useful to ensure that states and observables are encoded in a way which is both practical to implement and behaves favourably with respect to noise. In this section, we present such a notion of local encodings and state some basic properties; proofs are contained in Appendix A. A similar discussion is presented for the stronger case of local Hamiltonian encodings of Cubitt et al.<sup>11</sup>, and a discussion of the stability of local observable measurements to local noise is given by Trivedi et al.<sup>27</sup>.

**Definition 6** (Local state encoding). *Let  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$  and  $\mathcal{H}' = \otimes_{j=1}^{n'} \mathcal{H}'_j$ . We say that a state encoding  $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \rightarrow D(\mathcal{H}')$  is local if it has a Stinespring representation of the form*

$$\mathcal{E}_{\text{state}}(\rho) = \text{tr}_E[U(\rho \otimes |0\rangle\langle 0|_F)U^\dagger] , \quad (4)$$

where  $F = \otimes_k F_k$  and  $E = \otimes_l E_l$  are ancillary systems and  $U \in \text{U}(\mathcal{H} \otimes F, \mathcal{H}' \otimes E)$  is a constant-depth quantum circuit.

It is immediate that constant-depth quantum circuits (built from one-qubit and two-qubit gates) preserve locality. That is, given a local operator  $A$  on  $\mathcal{H} \otimes F$ , the operator  $UAU^\dagger$  is local (acting on the forward light cone of the support of  $A$ ) on  $\mathcal{H}' \otimes E$ , and similarly for the inverse  $U^\dagger$ . In fact, it is known in the theory of quantum cellular automata that this constraint is equivalent to representability as a constant-depth quantum circuit<sup>52</sup>.

For simulating physical systems, one particularly desirable feature of a simulator is local error back-propagation. That is, local noise on the simulator system should correspond in some way to local (perhaps realistic) noise on the target system. Ideally, we would like to prove that for any state  $\rho \in D(\mathcal{H})$  and local error channel  $\mathcal{N}' : D(\mathcal{H}') \rightarrow D(\mathcal{H}')$  on the simulator, there exists a corresponding local error channel  $\mathcal{N} : D(\mathcal{H}) \rightarrow D(\mathcal{H})$  on the target system satisfying

$$\mathcal{N}' \circ \mathcal{E}_{\text{state}}(\rho) \stackrel{?}{=} \mathcal{E}_{\text{state}} \circ \mathcal{N}(\rho) . \quad (5)$$

However, we cannot hope to prove this in general, since the noise operator  $\mathcal{N}'$  may take the simulator system outside the image of  $\mathcal{E}_{\text{state}}$ . Instead we have a slightly weaker version of this statement, which is a direct consequence of the causal structure of local state encodings.

**Proposition 7** (Local error back-propagation). *Let  $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \rightarrow D(\mathcal{H}')$  be a local state encoding as in Definition 6, and let  $\mathcal{N}' : D(\mathcal{H}') \rightarrow D(\mathcal{H}')$  be a channel whose Kraus operators  $\{X'_k\}$  each act on  $O(1)$  sites in  $\mathcal{H}'$ . Then there exists a channel  $\mathcal{N} : D(\mathcal{H} \otimes F) \rightarrow D(\mathcal{H} \otimes F)$  whose Kraus operators  $\{X_k\}$  each act on  $O(1)$  sites in  $\mathcal{H} \otimes F$ , and such that for all  $\rho \in \mathcal{H}$ ,*

$$\mathcal{N}' \circ \mathcal{E}_{\text{state}}(\rho) = \text{tr}_E[U\mathcal{N}(\rho \otimes |0\rangle\langle 0|_F)U^\dagger] . \quad (6)$$

In other words, local noise on the simulator corresponds to local noise on the target system and ancillary encoding system. The corresponding result with locality replaced by geometric locality holds in the case when the light cones of  $U$  are local with respect to the underlying geometry of the simulator and target systems.

Similarly, we have local forward-propagation under such an encoding, in the sense that local operations on a site  $\mathcal{H}_i$  to  $\rho \in D(\mathcal{H})$  will not affect the reduced density matrix  $\text{tr}_A[\mathcal{E}_{\text{state}}(\rho)]$ , where  $A$  is the forward light cone of  $\mathcal{H}_i$  under  $U$  in  $\mathcal{H}'$ .

We define local observable encodings analogously to the state encoding case.

**Definition 8** (Local observable encoding). *Let  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$  and  $\mathcal{H}' = \otimes_{j=1}^{n'} \mathcal{H}'_j$ . We say that an observable encoding  $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \rightarrow \text{Herm}(\mathcal{H}')$  is local if it is the adjoint (with respect to the Hilbert-Schmidt inner product) of a local state encoding.*

It is immediate from this definition that one can measure the encoded observable  $\mathcal{E}_{\text{obs}}(O)$  by first applying the constant-depth quantum circuit  $\mathcal{E}_{\text{obs}}^* : D(\mathcal{H}') \rightarrow D(\mathcal{H})$  to  $\rho' \in D(\mathcal{H}')$ , and then measuring  $O$ . When  $O$  is local, we can alternatively implement the measurement via a local POVM directly on the simulator system.

**Proposition 9** (Encoded measurements). *Let  $\mathcal{E}_{\text{obs}}$  be a local observable encoding as in Definition 8, and let  $O$  be a local operator on  $\mathcal{H}$ . Then  $\mathcal{E}_{\text{obs}}(O)$  can be measured using a local POVM on  $\mathcal{H}'$ .*

## Applications of the framework

In this section, we discuss some basic applications of our notion of analogue quantum simulation in the sense we have introduced in Definition 5. Firstly, we give an example of a trivial but illustrative situation in which encoding qudits into qubits incurs an unavoidable cost for low-energy encodings, but which is not an issue in our framework. We then demonstrate the robustness of the definition under noise, and show that it is consistent with the existing notion of simulation given in Definition 4. Finally, we note how Lieb-Robinson bounds can be used to reduce the overhead of simulating local observables.

**Qudits to qubits** To motivate this example, we first notice that the requirement of Cubitt et al.<sup>11</sup> (Definition 4) that the simulator Hamiltonian should reproduce the target dynamics in its low-energy subspace is too strong for some practical situations. As observed by the authors, this can require the simulator to use strong interactions to push unwanted states out of the low-energy subspace. Proposition 10 provides a formal statement of this fact (proved in Appendix B) in the context of encoding a simple qutrit Hamiltonian into qubits.

Here we consider qutrits with individual state spaces  $\mathbb{C}^3$  spanned by a basis  $\{|\downarrow\rangle, |0\rangle, |\uparrow\rangle\}$ . We write  $P_0^{(j)} = |0\rangle\langle 0|$  and  $P_{\uparrow}^{(j)} = |\uparrow\rangle\langle \uparrow|$ , where the superscript indicates that the projectors act on the  $j$ th qutrit.

**Proposition 10.** *Let  $\mathcal{H} = (\mathbb{C}^3)^{\otimes n}$  be the space of  $n$  qutrits acted on by the Hamiltonian*

$$H_n = \sum_{j=1}^n (P_0^{(j)} + P_{\uparrow}^{(j)}) . \quad (7)$$

*Suppose  $H'_n = \sum_{j=1}^K h'_j$  is a  $k$ -local Hamiltonian on  $\mathcal{H}' = (\mathbb{C}^2)^{\otimes m}$ , where  $m = O(n^{1+\alpha})$ , for  $\alpha \geq 0$  and  $k = O(1)$ . Assume the interaction hypergraph of  $H'_n$  has degree bounded by  $d = O(1)$ .*

If  $H'_n$  is a  $(\Delta, \eta, \epsilon)$ -simulation for  $H_n$  in the sense of Definition 4, for  $\eta \in [0, 1)$  and  $\epsilon \geq 0$ , then

$$\max_j \|h'_j\| = \Omega(n^{1-\alpha}(1 - \eta^2)) . \quad (8)$$

From (8) we see that simulating this simple system with a low-energy encoding, an interaction hypergraph of bounded degree, and bounded locality, requires either the qubit count or interaction energy (or a mixture) to scale unfeasibly with  $n$ . This constitutes a violation of the requirements of Definition 1 and imposes an unnecessary experimental requirement for the task of simulating non-interacting qutrits. The proof of this fact follows from a dimension-counting argument, since the state space of the qutrits cannot be surjectively encoded into the qubit simulator, see Fig. 1. In contrast, the simulation task is trivial in our framework given in Definition 5 because the low-energy encoding requirement is relaxed.

Letting  $H_n = \sum_{j=1}^n (P_0^{(j)} + P_\uparrow^{(j)})$  as in Proposition 10, we can simulate all observables under  $H_n$  on  $\mathcal{H}' = \otimes_{j=1}^n (\mathbb{C}^2 \otimes \mathbb{C}^2)$  via any isometry

$$V : \mathbb{C}^3 \rightarrow \mathbb{C}^2 \otimes \mathbb{C}^2 , \quad (9)$$

encoding each qutrit into two qubits. To realise a simulator in the sense of Definition 5, we let

$$\mathcal{E}_{\text{state}} : \rho \mapsto V^{\otimes n} \rho (V^{\otimes n})^\dagger , \quad \mathcal{E}_{\text{obs}} : O \mapsto V^{\otimes n} O (V^{\otimes n})^\dagger , \quad (10)$$

and

$$T_t = e^{-it\mathcal{E}_{\text{obs}}(H_n)}(\cdot)e^{it\mathcal{E}_{\text{obs}}(H_n)} , \quad (11)$$

which is just time evolution under a 2-local Hamiltonian with bounded strength interactions.

Although Proposition 10 does not necessarily rule out simulations in which the  $n$  qutrits are encoded into  $\Omega(n^2)$  qubits, such approaches suffer from a different problem. Generally, if each qudit in a  $D$ -dimensional system is encoded into  $\Omega(n^\alpha)$  qudits for  $\alpha > 0$ , whilst keeping the dimension fixed, then the inflated system size will necessarily cause the distances between encoded sites to grow with  $n$ . In a system of interacting qutrits (for which the proof of Proposition 10 still holds), this means that scaling interactions can be necessary to overcome Lieb-Robinson bounds and ensure that correlations can spread sufficiently fast through the enlarged system. The following simple geometric lemma provides some intuition for a quantitative lower bound on the growing length scales in such situations.

**Lemma 11.** *Let  $\{x_i\}_{i=1}^n$  be the points in a hypercube of side length  $L \sim n^{1/D}$  in the square lattice  $x_i \in \mathbb{Z}^D$ . Let  $\mathcal{E} : x_i \mapsto X_i \subseteq \mathbb{Z}^D$  be a map which encodes each point  $x_i$  into a connected set of points in  $\mathbb{Z}^D$  such that  $|X_i| = \Omega(n^\alpha)$  and  $X_i \cap X_j = \emptyset$ . Let  $d(x, y) : \mathbb{Z}^D \times \mathbb{Z}^D \rightarrow \mathbb{Z}$  be the taxicab metric on  $\mathbb{Z}^D$ .*

*For a radius  $R = O(L)$ , and any  $y \in \mathbb{Z}^D$ , the number of encoded points intersecting with the ball of radius  $R$  centred at  $y$  is bounded by*

$$|B_R(y)| := |\{X_i : \exists x \in X_i \text{ with } d(x, y) \leq R\}| = O(n^{1-\min\{\alpha, 1/D\}}) . \quad (12)$$

Letting  $\lambda = \min\{\alpha, 1/D\}$ , we see that there are at most  $O(n^{1-\lambda})$  sites  $X_j$  within radius  $R = O(L) = O(n^{1/D})$  of any  $X_i$ . On the other hand, there are at least  $\Omega(n^{1-\lambda})$  of the  $x_j$  within radius  $O(n^{(1-\lambda)/D})$  of  $x_i$  in the original lattice. In particular, this implies that there exist a pair of sites  $x_i, x_j$  with  $d(x_i, x_j) = O(n^{(1-\lambda)/D})$  whose encodings have  $d(X_i, X_j) = \Omega(n^{1/D})$  — in the encoded system, the distance is increased by a factor of  $n^{\lambda/D}$ .

The scalings here apply as a result of the requirement that analogue quantum simulators reproduce



the dynamics of a target system. In other situations, such as adiabatic quantum simulation in which an approximately simulated ground state is the only requirement, encodings with superlinear qubit overhead are possible<sup>53,54</sup>.

**Noisy analogue simulators** Suppose we have quantum channels  $T_t$ , for  $t \in [0, t_{\max}]$  which simulate some  $H \in \text{Herm}(\mathcal{H})$  with respect to  $\Omega_{\text{state}}$  and  $\Omega_{\text{obs}}$  up to accuracy  $\epsilon$  as in Definition 5, corresponding to encoding maps  $\mathcal{E}_{\text{state}}$  and  $\mathcal{E}_{\text{obs}}$ .

In practice, the experimental setup will suffer from noise in the steps of state preparation, evolution, and measurement. This will correspond to noisy versions of the above maps, which we denote by  $\tilde{T}_t$ ,  $\tilde{\mathcal{E}}_{\text{state}}$ , and  $\tilde{\mathcal{E}}_{\text{obs}}$ . For any  $O \in \Omega_{\text{obs}}$ ,  $\rho \in \Omega_{\text{state}}$ , we may bound the additional error in observable expectation values incurred by the noisy maps by

$$\begin{aligned}
& |\text{tr}[\mathcal{E}_{\text{obs}}(O)(T_t \circ \mathcal{E}_{\text{state}})(\rho)] - \text{tr}[\tilde{\mathcal{E}}_{\text{obs}}(O)(\tilde{T}_t \circ \tilde{\mathcal{E}}_{\text{state}}(\rho))]| \\
&= |\text{tr}[O(\mathcal{E}_{\text{obs}}^* \circ T_t \circ \mathcal{E}_{\text{state}} - \tilde{\mathcal{E}}_{\text{obs}}^* \circ \tilde{T}_t \circ \tilde{\mathcal{E}}_{\text{state}})(\rho)]| \\
&\leq \|\mathcal{E}_{\text{obs}}^* \circ T_t \circ \mathcal{E}_{\text{state}} - \tilde{\mathcal{E}}_{\text{obs}}^* \circ \tilde{T}_t \circ \tilde{\mathcal{E}}_{\text{state}}\|_{1 \rightarrow 1} \\
&\leq \|\mathcal{E}_{\text{obs}}^* - \tilde{\mathcal{E}}_{\text{obs}}^*\|_{1 \rightarrow 1} + \|T_t - \tilde{T}_t\|_{1 \rightarrow 1} + \|\mathcal{E}_{\text{state}} - \tilde{\mathcal{E}}_{\text{state}}\|_{1 \rightarrow 1} ,
\end{aligned} \tag{13}$$

where  $\|\cdot\|_{1 \rightarrow 1}$  denotes the one-to-one norm  $\|\Lambda\|_{1 \rightarrow 1} = \sup_{\rho} \|\Lambda(\rho)\|_1$  (defined as the induced trace norm<sup>55</sup> — note that this is in particular upper bounded by the diamond norm), and  $\mathcal{E}^*$  denotes the Hilbert-Schmidt dual of a superoperator  $\mathcal{E}$ . Hence the noisy simulator  $\tilde{T}_t$  also simulates  $H$  with respect to  $\Omega_{\text{state}}$  and  $\Omega_{\text{obs}}$ , up to error

$$\epsilon' \leq \epsilon + \sup_t \|T_t - \tilde{T}_t\|_{1 \rightarrow 1} + \|\mathcal{E}_{\text{state}} - \tilde{\mathcal{E}}_{\text{state}}\|_{1 \rightarrow 1} + \|\mathcal{E}_{\text{obs}}^* - \tilde{\mathcal{E}}_{\text{obs}}^*\|_{1 \rightarrow 1} . \tag{14}$$

**Local Hamiltonian simulation in a subspace** Suppose that  $H'$  is a  $(\Delta, \eta, \epsilon)$ -simulation of  $H$  as defined by Cubitt et al.<sup>11</sup> (Definition 4), corresponding to encodings  $\mathcal{E}_{\text{state}}$  and  $\mathcal{E}_{\text{obs}}$ , with the projector  $Q = 0$ . Here we show that the time evolution channel under  $H'$ ,  $(\cdot) \mapsto e^{-itH'}(\cdot)e^{itH'}$  gives a simulation in our sense, Definition 5.

We make use of the following lemmas. Lemma 12 ensures that measurement and time evolution are consistent with the encodings of Definition 4, and Lemma 13 bounds the error of  $(\Delta, \eta, \epsilon)$ -simulations under time evolution.

**Lemma 12** (Cubitt et al., Proposition 4<sup>11</sup>). *If  $\mathcal{E}_{\text{state}}$  and  $\mathcal{E}_{\text{obs}}$  are encodings as in Definition 4 and (2), then for all observables  $O$  and states  $\rho$  on the target system  $\mathcal{H}$ ,*

$$\text{tr}[\mathcal{E}_{\text{obs}}(O)\mathcal{E}_{\text{state}}(\rho)] = \text{tr}[O\rho] . \tag{15}$$

Moreover if the encoding is standard ( $\text{rank}(P) > 0$  in Definition 4) then

$$e^{-i\mathcal{E}_{\text{obs}}(H)t}\mathcal{E}_{\text{state}}(\rho)e^{i\mathcal{E}_{\text{obs}}(H)t} = \mathcal{E}_{\text{state}}(e^{-iHt}\rho e^{iHt}) . \tag{16}$$

**Lemma 13** (Cubitt et al., Proposition 28<sup>11</sup>). *Let  $H'$  be a  $(\Delta, \eta, \epsilon)$ -simulation of  $H$  in the sense of Definition 4 corresponding to encodings  $\mathcal{E}_{\text{obs}}$ ,  $\mathcal{E}_{\text{state}}$ . If  $\rho'$  is a state in the simulator system  $\mathcal{H}'$  satisfying*

$\mathcal{E}_{\text{obs}}(\mathbb{I})\rho' = \rho'$ , then for all  $t$

$$\|e^{-iH't}\rho'e^{iH't} - e^{-i\mathcal{E}_{\text{obs}}(H)t}\rho'e^{i\mathcal{E}_{\text{obs}}(H)t}\|_1 \leq 2\epsilon t + 4\eta. \quad (17)$$

Combining these lemmas, we see that for any observable  $O$  and state  $\rho$  on  $\mathcal{H}$ ,

$$\begin{aligned} & |\text{tr}[\mathcal{E}_{\text{obs}}(O)e^{-iH't}\mathcal{E}_{\text{state}}(\rho)e^{iH't}] - \text{tr}[Oe^{-iHt}\rho e^{iHt}]| \\ &= |\text{tr}[\mathcal{E}_{\text{obs}}(O)\left(e^{-iH't}\mathcal{E}_{\text{state}}(\rho)e^{iH't} - e^{-i\mathcal{E}_{\text{obs}}(H)t}\mathcal{E}_{\text{state}}(\rho)e^{i\mathcal{E}_{\text{obs}}(H)t}\right)]| \\ &\leq \|O\|(2\epsilon t + 4\eta). \end{aligned} \quad (18)$$

Hence the channels  $T_t : \rho' \mapsto e^{-iH't}\rho'e^{iH't}$ , for  $t \in [0, t_{\text{max}}]$  simulate  $H$  in the sense of Definition 5 with respect to any  $\Omega_{\text{state}}$  and  $\Omega_{\text{obs}}$ , up to error

$$\epsilon' \leq 2\epsilon t_{\text{max}} + 4\eta. \quad (19)$$

This provides some consistency between existing work and our notion of simulation; we have shown that evolution under a simulator Hamiltonian in the sense of Cubitt et al.<sup>11</sup> constitutes an analogue quantum simulator in our framework given by Definition 5.

### Short-time simulation with Lieb-Robinson bounds

One advantage of only requiring the simulation of a particular set of observables  $\Omega_{\text{obs}}$  in Definition 5, as opposed to reproducing the entire physical system, is that one can take advantage of the limited spread of correlations for short-time dynamics<sup>24</sup>. The idea of exploiting Lieb-Robinson bounds to reduce necessary hardware overhead has already been considered for the study of many-body quantum states on quantum computers<sup>25,26</sup>, and more recently in the setting of analogue simulators<sup>27</sup>. We explain here how the latter fits into our framework.

Consider the case of a Hamiltonian  $H_n$  on a  $d$ -dimensional lattice of  $n$  qubits  $\mathcal{H} \cong (\mathbb{C}^2)^{\otimes n}$ , such that

$$H_n = \sum_{x=1}^n h_x, \quad (20)$$

where the  $h_x$  is a nearest-neighbour local interaction with  $\|h_x\| \leq 1$ , translated to position  $x$  in the lattice, so that  $H_n$  is translationally invariant.

If one is only interested in simulating the finite-time dynamics of a few local observables  $\Omega_{\text{obs}}$  which are contained within a small neighbourhood of the origin, starting from a state  $\rho = |0\rangle\langle 0|^{\otimes n}$ , then it is sufficient (up to exponentially small error) to simulate a far smaller subsystem, corresponding to the Lieb-Robinson light cone, as in Fig. 7. This situation is studied by Trivedi et al.<sup>27</sup>, in particular for the thermodynamic limit  $n \rightarrow \infty$ .

Let  $H_m = \sum_{y=1}^m h_y$  be the simulator Hamiltonian, defined identically to  $H_n$  but on a lattice of size  $m < n$ ,  $\mathcal{H}' \cong (\mathbb{C}^2)^{\otimes m}$ . We encode  $\rho$  and  $O$  simply by restricting them to the smaller subsystem. Then a simulation of an observable  $O \in \Omega_{\text{obs}}$  up to accuracy  $\epsilon$ , satisfying

$$|\text{tr}[Oe^{-iHnt}\rho e^{iHnt}] - \text{tr}[\mathcal{E}_{\text{obs}}(O)e^{-iH_mt}\mathcal{E}_{\text{state}}(\rho)e^{iH_mt}]| \leq \epsilon, \quad (21)$$

can be accomplished in the large- $n$  regime for all  $t \in [0, t_{\text{max}}]$  if one takes  $m = O(\log^d(1/\epsilon) + t_{\text{max}}^d)$  (see Trivedi et al.<sup>27</sup>, Lemma 1).

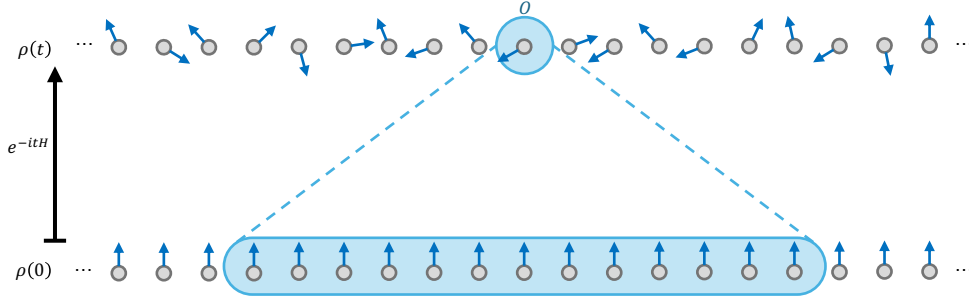


Figure 7: Simulation with Lieb-Robinson bounds. Simulation of a 1-dimensional spin system under a Hamiltonian  $H$  for time  $t$ . In theory, the system extends infinitely, but to estimate the value of a local observable  $O$  it is only necessary to simulate a subsystem corresponding to the Lieb-Robinson light cone.

### Modular encodings and gadgets

In this section, we focus on the case of a simulator channel  $T_t$  given by time evolution under a local simulator Hamiltonian  $H'$ , which should reproduce the dynamics of the local target Hamiltonian  $H = \sum_i H_i$ . In light of the size-independence requirement of Definition 1, it is natural to encode each  $H_i$  term separately into some term  $H'_i$ , but systematically doing so is a non-trivial task: we need the encoded terms to interact with each other in a way which mimics the original system.

This problem can be tackled using perturbative gadgets. Perturbative gadgets were initially introduced by Kempe et al.<sup>42</sup> as a means of proving QMA-completeness of the 2-local Hamiltonian problem by reduction from the 3-local case<sup>30</sup>, and have since been used extensively in the field of Hamiltonian complexity theory. In this work, we especially focus on the use of gadgets for Hamiltonian locality reduction, though it should be noted that perturbative gadgets can also be used to simplify the structure of the interaction hypergraph<sup>17</sup> and in general to reduce Hamiltonians to more restrictive families of interactions<sup>56,57,44</sup>. Moreover, beyond Hamiltonian complexity-theoretic results, gadgets can be tailored to improve the performance of variational quantum algorithms<sup>34</sup>.

In this work, we introduce a formalism which we argue encompasses any attempt at gadgetisation, in a sense which we make precise (Definition 14), in order to prove general properties of such constructions. Note that our approach, and the  $(\eta, \epsilon)$  accuracy parameters, are closely related to those used in other definitions of simulation<sup>11,18</sup>. We refine the approach of the latter by generalising to a potentially non-perturbative regime and by considering the feature of combining well with other interactions as a generic requirement for gadgets. We use these results to argue that any size-independent encoding of a Hamiltonian  $H$  into another  $H'$  cannot reduce the locality of interactions (for example, reducing a 3-local Hamiltonian to a 2-local Hamiltonian).

The setup is as follows: we consider a large system  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$ , within which a local interaction  $H \in \text{Herm}(\mathcal{H})$  acts on a subsystem of  $O(1)$  sites. With the introduction of a small ancillary system  $\mathcal{A}$ , we aim to replace  $H$  by some gadget  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$ , which acts on  $O(1)$  sites in  $\mathcal{H}$  and  $\mathcal{A}$ .

A simulator Hamiltonian in the sense of Definition 5 need not necessarily capture the entire spectrum of its target Hamiltonian. In this case, however, we are thinking of  $H$  as a single interaction in a larger system, and as such we cannot generally assume that its eigenspaces will be preserved under

time evolution. Therefore, we require as a minimum that  $H'$  should (when restricted to some subspace defined by a projector  $P'$ ) approximately reproduce the full spectrum of  $H$ . Moreover, for  $H'$  to be a useful gadget, it must combine well with other Hamiltonian terms acting on  $\mathcal{H}$ . That is to say, there should exist  $P' \in \text{Proj}(\mathcal{H} \otimes \mathcal{A})$  such that  $P'(H' + H_{\text{else}} \otimes \text{I})P'$  approximates the spectrum of  $H + H_{\text{else}}$ , for any  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$  (see Fig. 3(a)). We formalise this with the following definition.

**Definition 14** ( $(\zeta, \epsilon)$ -gadget property). *Given a Hamiltonian  $H \in \text{Herm}(\mathcal{H})$  acting on a system  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$ , and  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$  for  $\mathcal{A}$  an ancillary system, we say that  $(H', \mathcal{A})$  satisfies the  $(\zeta, \epsilon)$ -gadget property for  $H$  if there exists  $P' \in \text{Proj}(\mathcal{H} \otimes \mathcal{A})$ ,  $\tilde{P} \in \text{Proj}(\mathcal{A}) \setminus \{0\}$  such that, for any  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$ , there exists a unitary  $\tilde{U}_{H_{\text{else}}} \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  with*

$$\|P'(H' + H_{\text{else}} \otimes \text{I})P' - \tilde{U}_{H_{\text{else}}}\left((H + H_{\text{else}}) \otimes \tilde{P}\right)\tilde{U}_{H_{\text{else}}}^\dagger\| \leq \epsilon + \zeta\|H_{\text{else}}\|. \quad (22)$$

In other words,  $(H', \mathcal{A})$  satisfies the  $(\zeta, \epsilon)$ -gadget property for  $H$  if, when restricted a subspace defined by  $P'$ ,  $H' + H_{\text{else}} \otimes \text{I}$  approximates the spectrum of  $H + H_{\text{else}}$  up to error  $\epsilon + \zeta\|H_{\text{else}}\|$ . Notice that  $\tilde{P}$  is almost arbitrary; its rank determines the multiplicity of each eigenvalue of  $H + H_{\text{else}}$  in the simulator system, but otherwise it can be rotated by  $\tilde{U}_{H_{\text{else}}}$ , which rotates the eigenvectors of  $(H + H_{\text{else}}) \otimes \tilde{P}$  approximately onto those of  $P'(H' + H_{\text{else}} \otimes \text{I})P'$ .

As noted by Cubitt et al.<sup>11</sup>, there are two distinct types of gadgets used in literature:

- Mediator gadgets, in which ancillary qubits are inserted between logical qubits to mediate interactions, and
- Subspace gadgets, in which single logical qubits are encoded into several physical qubits, restricted to a two-dimensional subspace by strong interactions.

Definition 14 encompasses the former, but not the latter. Qualitatively this is because whereas mediator gadgets replace interactions, subspace gadgets replace entire qubits, including all of the interactions they take part in. It would be possible to extend our formalism to subspace gadgets, by restricting the range of  $H_{\text{else}}$  in Definition 14 to terms which do not interact with the target qubit. We do not consider this here, however, for brevity and because subspace gadgets do not reduce the locality of interactions, which is our primary motivation for this section.

Although Definition 14 is a natural requirement, it is not convenient to work with due to the appearance of the general  $H_{\text{else}}$  acting on the entire of  $\mathcal{H}$ , upon which  $\tilde{U}$  depends. The following alternative definition does not suffer from this problem.

**Definition 15** ( $(\eta, \epsilon)$ -gadget). *Let  $H \in \text{Herm}(\mathcal{H})$  be a Hamiltonian on a Hilbert space  $\mathcal{H}$ , and let  $\mathcal{A}$  be an ancillary Hilbert space. For  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$ , we say that  $(H', \mathcal{A})$  is a  $(\eta, \epsilon)$ -gadget for  $H$  if there exists  $P \in \text{Proj}(\mathcal{A}) \setminus \{0\}$  and  $U \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  such that*

$$\|U - \text{I}\| \leq \eta, \quad \|P'H'P' - U(H \otimes P)U^\dagger\| \leq \epsilon, \quad (23)$$

where  $P' = U(\text{I} \otimes P)U^\dagger \in \text{Proj}(\mathcal{H} \otimes \mathcal{A})$ .

The advantage of Definition 15 is that it is stated in terms of a local rather than global property. Assuming that  $H, H', P'$  act on only  $O(1)$  sites in  $\mathcal{H}$  and  $\mathcal{A}$ , we can without loss of generality restrict to this significantly smaller subspace to check whether  $H'$  is a gadget. This is in contrast with Definition 14,

which requires us to in principle consider interactions over the full  $n$ -site space in order to check the gadget property.

To motivate the use of Definition 15, we show that the above notions are in correspondence; things that look like gadgets are always gadgets, and vice-versa. This is formalised by the following two theorems, proved in Appendix C.

**Theorem 16** ( $(\eta, \epsilon)$ -gadgets have the  $(\zeta, \epsilon)$ -gadget property). *Suppose that  $(H', \mathcal{A})$  is a  $(\eta, \epsilon)$ -gadget for  $H$ . Then  $(H', \mathcal{A})$  satisfies the  $(\zeta, \epsilon)$ -gadget property for  $H$ , where  $\zeta = O(\eta)$ .*

**Theorem 17** (The  $(\zeta, \epsilon)$ -gadget property requires a  $(\eta, \epsilon)$ -gadget). *Suppose that  $(H', \mathcal{A})$  satisfies the  $(\zeta, \epsilon)$ -gadget property for  $H$ , where  $H, H'$ , and  $P'$  act on  $O(1)$  sites in  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$ . Then  $(H', \mathcal{A})$  is a  $(\eta, \epsilon')$ -gadget for  $H$ , where  $\eta = O(\epsilon) + O(\zeta^{\frac{1}{2}})$  and  $\epsilon' = O(\epsilon) + O(\zeta)$ .*

The roles of the  $\eta$  and  $\epsilon$  parameters are to bound the error in the eigenvectors and eigenvalues respectively. Roughly speaking,  $\eta$  quantifies how well the gadget combines with other terms, and  $\epsilon$  quantifies the spectral error of the gadget in isolation. A good gadget requires both of these parameters to be small. In the next section we present a 3-to-2 local gadget which is an extreme case of this, with  $\epsilon = 0$  at the cost of a large  $\eta$  error.

Prior work in Hamiltonian complexity theory has focused on gadgetisation in the context of ground state estimation<sup>30,44,18</sup> or simulation in a low energy subspace<sup>11</sup>; as a result, a case of particular relevance is when  $P'$  projects onto the low-energy subspace of  $H'$ . For  $\Delta \in \mathbb{R}$ , we write  $P_{\leq \Delta(H')}$  for the projector onto the span of the eigenvectors of  $H'$  with eigenvalues in the range  $(-\infty, \Delta]$ .

**Definition 18** ( $(\Delta, \eta, \epsilon)$ -gadget). *Let  $H \in \text{Herm}(\mathcal{H})$  be a Hamiltonian on a Hilbert space  $\mathcal{H}$ , and let  $\mathcal{A}$  be an ancillary Hilbert space. For  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$ , we say that  $(H', \mathcal{A})$  is a  $(\Delta, \eta, \epsilon)$ -gadget for  $H$  if there exists  $P \in \text{Proj}(\mathcal{A}) \setminus \{0\}$ , and  $U \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  such that  $P_{\leq \Delta(H')} = U(\text{I} \otimes P)U^\dagger$ , and*

$$\|U - \text{I}\| \leq \eta, \quad \|P_{\leq \Delta(H')} H' P_{\leq \Delta(H')} - U(H \otimes P)U^\dagger\| \leq \epsilon. \quad (24)$$

*In other words, the pair  $(H', \mathcal{A})$  satisfy Definition 15, in the special case where we can use  $P' = P_{\leq \Delta(H')}$ .*

Notice that Definition 18 imposes a significantly stronger requirement on  $H'$  than Definition 15; a priori there is no reason to expect that there will exist any choice of  $P$  and  $U$  such that  $P_{\leq \Delta(H')} = U(\text{I} \otimes P)U^\dagger$ . Definitions 15 and 18 are sufficient to guarantee desirable combination properties, and are satisfied by widely-used constructions.

## Examples of gadgets

Lemmas 4-7 of Bravyi et al.<sup>18</sup> can be naturally adapted to give several constructions for  $(\Delta, \eta, \epsilon)$  gadgets, which we use to demonstrate that Definition 15 encompasses commonly-used techniques. In the following we take  $\mathcal{H}' = \mathcal{H} \otimes \mathcal{A}$ , and  $\mathcal{A} \cong \mathbb{C}^2$ . For  $V$  an operator on  $\mathcal{H}'$  we write it in block-diagonal form with respect to the basis of  $\mathcal{A}$  as

$$V = \begin{pmatrix} V_{00} & V_{01} \\ V_{10} & V_{11} \end{pmatrix}, \quad (25)$$

where, for instance,  $V_{00} = (\text{I} \otimes \langle 0|)V(\text{I} \otimes |0\rangle)$ .

**Lemma 19** (First-order gadgets, adapted from Bravyi et al.<sup>18</sup>). *Suppose  $H \in \text{Herm}(\mathcal{H})$  and  $V \in \text{Herm}(\mathcal{H}')$  are such that*

$$\|H - V_{00}\| \leq \frac{\epsilon}{2}. \quad (26)$$

*Then  $H' = \Delta H_0 + V$  defines a  $(O(\Delta), \eta, \epsilon)$ -gadget for  $H$ , where  $H_0 = \mathbb{I} \otimes |1\rangle\langle 1|$ , provided that  $\Delta \geq O(\epsilon^{-1}\|V\|^2 + \eta^{-1}\|V\|)$ .*

**Lemma 20** (Second-order gadgets, adapted from Bravyi et al.<sup>18</sup>). *Let  $H \in \text{Herm}(\mathcal{H})$ , and suppose  $V^{(1)}, V^{(0)} \in \text{Herm}(\mathcal{H}')$  are such that  $\|V^{(1)}\|, \|V^{(0)}\| \leq \Lambda$ ,  $V_{10}^{(0)} = V_{01}^{(0)} = V_{00}^{(1)} = 0$ , and*

$$\|H - V_{00}^{(0)} + V_{01}^{(1)}V_{10}^{(1)}\| \leq \frac{\epsilon}{2}. \quad (27)$$

*Then  $H' = \Delta H_0 + \Delta^{\frac{1}{2}}V^{(1)} + V^{(0)}$  is a  $(O(\Delta), \eta, \epsilon)$ -gadget for  $H$ , where  $H_0 = \mathbb{I} \otimes |1\rangle\langle 1|$ , if*

$$\Delta \geq O(\epsilon^{-2}\Lambda^6 + \eta^{-2}\Lambda^2). \quad (28)$$

**Lemma 21** (Third-order gadgets, adapted from Bravyi et al.<sup>18</sup>). *Let  $H \in \text{Herm}(\mathcal{H})$ , and suppose  $V^{(2)}, V^{(1)}, V^{(0)} \in \text{Herm}(\mathcal{H}')$  are such that  $\|V^{(2)}\|, \|V^{(1)}\|, \|V^{(0)}\| \leq \Lambda$ ,  $V_{10}^{(1)} = V_{01}^{(1)} = V_{10}^{(0)} = V_{01}^{(0)} = 0$ ,  $V_{00}^{(2)} = 0$ ,*

$$\|H - V_{00}^{(0)} - V_{01}^{(2)}V_{11}^{(2)}V_{10}^{(2)}\| \leq \frac{\epsilon}{2}, \quad \text{and} \quad V_{00}^{(1)} = V_{01}^{(2)}V_{10}^{(2)}. \quad (29)$$

*Then  $H' = \Delta H_0 + \Delta^{\frac{2}{3}}V^{(2)} + \Delta^{\frac{1}{3}}V^{(1)} + V^{(0)}$  is a  $(O(\Delta), \eta, \epsilon)$ -gadget for  $H$ , where  $H_0 = \mathbb{I} \otimes |1\rangle\langle 1|$ , if*

$$\Delta \geq O(\epsilon^{-3}\Lambda^{12} + \eta^{-3}\Lambda^3). \quad (30)$$

We illustrate the application of these lemmas to our definition with the following ubiquitous gadgets from Oliviera et al.<sup>17</sup>:

Given a target Hamiltonian  $H = A \otimes B \in \text{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)$ , the subdivision gadget on  $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C$  (where  $\mathcal{H}_C \cong \mathbb{C}^2$ ) is defined by

$$H' = \Delta H_0 + \Delta^{\frac{1}{2}}V^{(1)} + V^{(0)}, \quad (31)$$

where

$$H_0 = \mathbb{I} \otimes \mathbb{I} \otimes |1\rangle\langle 1|, \quad (32)$$

$$V^{(1)} = \frac{1}{\sqrt{2}}(-A \otimes \mathbb{I} + \mathbb{I} \otimes B) \otimes X, \quad (33)$$

$$V^{(0)} = \frac{1}{2}(A^2 \otimes \mathbb{I} + \mathbb{I} \otimes B^2) \otimes \mathbb{I}. \quad (34)$$

Then by Lemma 20 we see that, for sufficiently large  $\Delta$ ,  $(H', \mathcal{H}_C)$  defines a  $(O(\Delta), \eta, \epsilon)$ -gadget for  $H$  (see Fig. 8(a)).

Given a target Hamiltonian  $H = A \otimes B \otimes C \in \text{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$ , the 3-to-2 local gadget on  $\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C \otimes \mathcal{H}_D$  (where  $\mathcal{H}_D \cong \mathbb{C}^2$ ) is defined by

$$H' = \Delta H_0 + \Delta^{\frac{2}{3}}V^{(2)} + \Delta^{\frac{1}{3}}V^{(1)} + V^{(0)}, \quad (35)$$

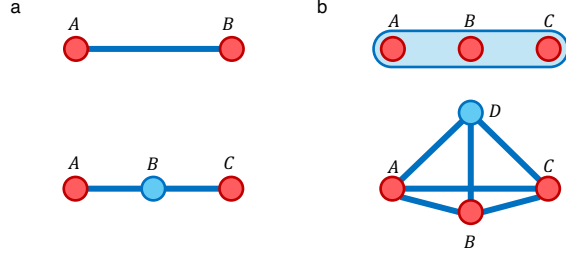


Figure 8: Existing gadgets. (a) Interaction hypergraphs of a 2-system interaction before (above) and after (below) the use of the subdivision gadget. (b) Interaction hypergraphs of a 3-system interaction before (above) and after (below) the use of the 3-to-2 gadget.

where

$$H_0 = \mathbb{I} \otimes \mathbb{I} \otimes \mathbb{I} \otimes |1\rangle\langle 1| , \quad (36)$$

$$V^{(2)} = \frac{1}{\sqrt{2}}(-A \otimes \mathbb{I} + \mathbb{I} \otimes B) \otimes \mathbb{I} \otimes X - \mathbb{I} \otimes \mathbb{I} \otimes C \otimes |1\rangle\langle 1| , \quad (37)$$

$$V^{(1)} = \frac{1}{2}(-A \otimes \mathbb{I} + \mathbb{I} \otimes B)^2 \otimes \mathbb{I} \otimes \mathbb{I} , \quad (38)$$

$$V^{(0)} = \frac{1}{2}(A^2 \otimes \mathbb{I} + \mathbb{I} \otimes B^2) \otimes C \otimes \mathbb{I} . \quad (39)$$

By Lemma 21 we see that, for sufficiently large  $\Delta$ ,  $(H', \mathcal{H}_D)$  defines a  $(O(\Delta), \eta, \epsilon)$ -gadget for  $H$  (see Fig. 8(b)).

We provide the following example to illustrate the importance of the  $\eta$  parameter as a quantifier of how well a gadget combines with other terms.

Let  $H = A \otimes B \otimes C \in \text{Herm}((\mathbb{C}^2)^{\otimes 3})$  be a 3-qubit interaction, and diagonalise  $A$ ,  $B$ , and  $C$  as

$$A = \lambda_0^A |0\rangle\langle 0| + \lambda_1^A |1\rangle\langle 1| , \quad B = \lambda_0^B |0\rangle\langle 0| + \lambda_1^B |1\rangle\langle 1| , \quad C = \lambda_0^C |0\rangle\langle 0| + \lambda_1^C |1\rangle\langle 1| . \quad (40)$$

Let  $H' \in \text{Herm}((\mathbb{C}^2)^{\otimes 4})$  be defined as

$$\begin{aligned} H' &= \lambda_0^B (A - \lambda_0^A \mathbb{I}) \otimes \mathbb{I} \otimes \mathbb{I} \otimes C \\ &\quad + \lambda_1^B \mathbb{I} \otimes (A - \lambda_0^A \mathbb{I}) \otimes \mathbb{I} \otimes C \\ &\quad + \lambda_0^A \mathbb{I} \otimes \mathbb{I} \otimes B \otimes C , \end{aligned} \quad (41)$$

and let  $P' \in \text{Proj}((\mathbb{C}^2)^{\otimes 4})$  be

$$P' = (\mathbb{I} \otimes |0\rangle\langle 0| \otimes |0\rangle\langle 0| + |0\rangle\langle 0| \otimes \mathbb{I} \otimes |1\rangle\langle 1|) \otimes \mathbb{I} . \quad (42)$$

Then in fact the restriction of  $H'$  to the image of  $P'$  exactly reproduces the spectrum of  $H$ . This hence defines a 3-to-2  $(\eta, 0)$ -gadget — or a  $(\Delta, \eta, 0)$ -gadget, if one adds a term of the form  $O(\Delta)(\mathbb{I} - P')$  to  $H'$ . The caveat is that this gadget has a large  $\eta$  parameter, and hence it does not combine well with other interactions. For instance, in Definition 15 we might take  $P = |0\rangle\langle 0| \otimes \mathbb{I} \otimes \mathbb{I}$ , and  $U = (\mathbb{F} \otimes |0\rangle\langle 0| + \mathbb{I} \otimes \mathbb{I} \otimes |1\rangle\langle 1|) \otimes \mathbb{I}$ , where  $\mathbb{F}$  is the two-qubit swapping operator. This gives  $\eta = 2$ .

The construction of  $H'$  can be thought of as splitting the  $A$  qubit into two qubits (see Fig. 9), and

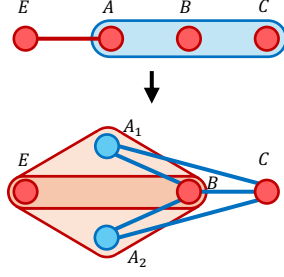


Figure 9: The exact 3-to-2 gadget. The (blue) 3-local interaction between  $A$ ,  $B$ , and  $C$  is replaced by a series of (blue) 2-local interactions, where the  $A$  site has been split into two sites  $A_1$  and  $A_2$ . However after this process, the 2-local interaction (red) between  $A$  and another qubit  $E$  is replaced by two 3-local interactions between  $E, A_1, B$  and  $E, A_2, B$ . Compare this with Fig. 8(b), for which additional interactions on qubit  $A$  will remain on qubit  $A$  of the gadgetised Hamiltonian without any need for adjustment.

controlling whether the first or second qubit is excited depending on the value of the  $B$  qubit. Therefore, if the full Hamiltonian contains another interaction term which acts on the  $A$  site in  $H$ , then the locality of this term will be increased under the gadgetisation procedure. Such a gadget cannot be used to systematically reduce the locality of a Hamiltonian with many interactions.

### Gadget combination results

The following results show that gadgets satisfying Definition 15 or Definition 18 can be systematically combined as desired. Our techniques and proofs extend prior work<sup>31,17,58</sup>, using the convenient formalism of the direct rotation<sup>59</sup>. The scalings of the parameters  $\eta', \epsilon'$  are not necessarily optimal, though they sufficient for application to the subdivision and 3-to-2 gadget constructions exhibited above. The proofs of our gadget combination results can be found in Appendix D.

We summarise the setup below, which will be used throughout the following results.

**Setup 22.** Let  $H \in \text{Herm}(\mathcal{H})$  be a Hamiltonian on  $n$  sites,  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$ . Assume  $H = \sum_{i=1}^N H_i$ , where  $N = O(n)$ , such that each  $H_i$  acts on at most  $k = O(1)$  of the sites  $\mathcal{H}_i$ , and each site participates in at most  $d = O(1)$  interactions. Assume also that  $H$  has bounded interaction strengths, that is,  $\|H_i\| \leq J$  for all  $i$ .

In the below propositions we consider a family (depending on  $n$ ) of gadgets  $(H'_i, \mathcal{A}_i)$  for  $H_i$ , with  $U_i, P_i$ , and  $P'_i$  defined as in Definition 15, for each  $i$ . Assume that  $\mathcal{A}_i$  consists of  $O(1)$  ancillary sites and that  $H'_i$  is a local Hamiltonian consisting of  $O(1)$  interactions, such that

$$\|H'_i\| \leq J' , \quad \|(\mathbb{I} \otimes P_i) H'_i (\mathbb{I} \otimes P_i^\perp)\| \leq J'_O . \quad (43)$$

Firstly, we state the main result: that gadgets as in Definition 15 may be systematically combined to produce new gadgets.

**Proposition 23** (Parallel  $(\eta, \epsilon)$ -gadget combination). Let  $H = \sum_i H_i$  be as in Setup 22, and suppose that each  $(H'_i, \mathcal{A}_i)$  defines a  $(\eta, \epsilon)$ -gadget for  $H_i$ .



Define

$$H' = \sum_i H'_i \in \text{Herm}(\mathcal{H} \otimes (\otimes_i \mathcal{A}_i)) . \quad (44)$$

Then  $(H', \otimes_i \mathcal{A}_i)$  is a  $(\eta', \epsilon')$ -gadget for  $H$ , where

$$\epsilon' = O(n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J') , \quad \eta' = O(n\eta) . \quad (45)$$

For completeness, we also prove a similar result that  $(\Delta, \eta, \epsilon)$ -gadgets can be combined to create a new  $(\Delta', \eta', \epsilon')$ -gadget. It follows from Proposition 23 that the combination of many  $(\Delta, \eta, \epsilon)$ -gadgets defines a  $(\eta', \epsilon')$ -gadget, however it still remains to show that the projector  $P'$  in the sense of Definition 15 may be taken as a low-energy projector  $P_{\leq \Delta'}(H')$ .

**Proposition 24** (Parallel  $(\Delta, \eta, \epsilon)$ -gadget combination). *Let  $H = \sum_i H_i$  be as in Setup 22, and suppose that each  $(H'_i, \mathcal{A}_i)$  defines a  $(\Delta, \eta, \epsilon)$ -gadget for  $H_i$ , where*

$$\Delta \geq \frac{\|H\| + J + N(\epsilon + 2J\eta)}{\frac{1}{4} - 2\eta} = O(nJ) , \quad (46)$$

and assume that the scaling of  $\eta$  with  $n$  is bounded as

$$\eta = o(n^{-\frac{1}{2}}) , \quad (47)$$

and moreover that, for large  $J'$ ,

$$n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J' = o(J') , \quad J' = O(\Delta) . \quad (48)$$

Define

$$H' = \sum_i H'_i \in \text{Herm}(\mathcal{H} \otimes (\otimes_i \mathcal{A}_i)) . \quad (49)$$

Then  $(H', \otimes_i \mathcal{A}_i)$  is a  $(\Delta', \eta', \epsilon')$ -gadget for  $H$ , where

$$\Delta' = \frac{1}{2}\Delta , \quad \epsilon' = O(n\epsilon + n\eta J + n\eta^3 J'_O + n^3 \eta^4 J') , \quad \eta' = O(n\eta) . \quad (50)$$

For an example of how these conditions can be satisfied, consider the case of combining many of the 3-to-2 gadgets described above. Setting  $J = 1$  for convenience, we have  $J' = \Theta(\Delta)$ ,  $J'_O = \Theta(\Delta^{2/3})$ , and  $\epsilon, \eta = O(\Delta^{-1/3})$ . The errors  $\epsilon'$  and  $\eta'$  both grow as  $O(n\Delta^{-1/3})$ , so a good gadget will require  $\Delta = \Omega(n^3)$ . A direct computation verifies that this condition also ensures that (46-48) are satisfied. Hence reduction from a 3-local to 2-local Hamiltonian in this way requires interaction strengths to scale as  $n^3$ .

To combine  $(\Delta, \eta, \epsilon)$  gadgets using Proposition 24 requires the unappealing conditions of (46)-(47), which explicitly require the gadget energies to scale with  $n$ . In fact, as noted by Bravyi et al.<sup>31</sup>, the regime of bounded-strength interactions does still allow approximation of the ground state energy of  $H$  — the caveat being that the errors are extensive. Below is a generalisation of their main result.

**Theorem 25** (Ground state energy estimation with  $(\Delta, \eta, \epsilon)$ -gadgets, generalising Bravyi et al., Theorem 1<sup>31</sup>). *Let  $H = \sum_i H_i$  be as in Setup 22, and suppose that each  $(H'_i, \mathcal{A}_i)$  defines a  $(\Delta, \eta, \epsilon)$ -gadget for  $H_i$ .*

Define

$$H' = \sum_i H'_i \in \text{Herm}(\mathcal{H} \otimes (\otimes_i \mathcal{A}_i)) . \quad (51)$$

Then the ground state energies of  $H$  and  $H'$  satisfy

$$|\lambda_0(H) - \lambda_0(H')| = O(n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J') . \quad (52)$$

## Gadget energy scaling

Here we present the main result of the section: general locality reduction gadgets cannot exist without unfavourably scaling energies. This result holds in the most general setting of  $(\eta, \epsilon)$ -gadgets (Definition 15), and hence follows even from the relaxed  $(\zeta, \epsilon)$ -gadget property of Definition 14.

**Theorem 26** (Gadget energy scaling). *Let  $\mathcal{H} = (\mathbb{C}^2)^{\otimes k}$  be the space of  $k = O(1)$  qubits, and let  $H$  be the  $k$ -fold tensor product of Pauli  $Z$  operators with strength  $J > 0$ ,*

$$H = J \bigotimes_{i=1}^k Z_i . \quad (53)$$

*Suppose  $(H', \mathcal{A})$  is a  $(\eta, \epsilon)$ -gadget for  $H$  for  $H'$  a  $k'$ -local Hamiltonian, where  $k' < k$ .*

*Then, provided  $\epsilon < J$ , the gadget must have energy scale  $\|H'\| \geq \frac{J-\epsilon}{\eta} = \Omega(\eta^{-1})$ .*

The method of proof (found in Appendix E) is simple, and very likely does not provide an optimal lower bound for  $\|H'\|$ , due to the lack of any dependence on  $k$ . We expect that such dependence should be present; any approach which iteratively lowers the locality of an interaction from  $k$ -local to 2-local will accumulate scalings from each round of gadgetisation, but this does not rule out a more direct approach. Existing methods to reduce locality, such as the subdivision and 3-to-2 gadgets of Oliveira et al.<sup>17</sup> and higher-order gadgets<sup>60,34</sup>, give scalings that suggest that any  $k$ -to-2-local gadget construction should require energies which scale exponentially in  $k$ . The question of whether such exponential scaling is the best possible was first raised by Bravyi et al.<sup>31</sup>, and is still unresolved. Using the formalism introduced here, this problem can be precisely stated, and optimisation of Theorem 26 may provide a negative result. Furthermore, we expect that it may be possible to answer similar questions about gadget energy scaling in other cases, for example in simplifying the structure of an interaction graph or reducing to smaller families of interactions.

The significance of Theorem 26 is that it essentially rules out a size-independent (Definition 1) simulation of a  $k$ -local Hamiltonian  $H$  by another  $k'$ -local Hamiltonian  $H'$  for  $k' < k$ , for the following reason. Any modular encodings require the use of term-by-term gadgets, which must each satisfy the  $(\zeta, \epsilon)$ -gadget property (Definition 14) with  $\zeta, \eta = O(n^{-1})$  to guarantee that they can be combined (since the rest of the Hamiltonian will have  $\|H_{\text{else}}\| = O(n)$ ). By Theorem 17, this requires the use of  $(\eta, \epsilon)$ -gadgets (Definition 15) with  $\eta = O(n^{-1/2})$ , and by Theorem 26 this will require interactions which scale at least as  $\Omega(n^{1/2})$ .

A couple of notes on gadget energy scalings in existing work: Bausch<sup>61</sup> gives a method to reduce the exponential or doubly-exponential scaling in perturbative Hamiltonians to polynomial scaling, and Cao et al.<sup>33</sup> present gadgets whose interaction strengths do not grow with accuracy. However, both cases violate size-independence (Definition 1) in other ways such as polynomial scaling in the number of simulator qubits or instead shrinking the interaction strengths.

## Gadgets from the quantum Zeno effect

In this section, we demonstrate an alternative approach for reducing the locality of an interaction in a Hamiltonian — a task for which Theorem 26 establishes the need for energies which scale with the size of the system, when conventional gadgets are used. The construction presented here, however, uses the freedom afforded by the general simulation channel  $T_t$  in Definition 5 to take advantage of an additional resource: dissipation.

We will see that, despite some impractical features for experimental implementation, this approach offers a theoretical improvement in scalings over the conventional gadget techniques discussed earlier in the section. Additionally, this construction captures a key feature of our framework for analogue simulators given in Definition 5 in contrast with existing work: we define simulators in terms of their dynamic behaviour, rather than in terms of the properties of static Hamiltonians.

For the process we describe here, we repeatedly refer to measurement for conceptual simplicity when talking about probabilities, but this terminology is somewhat misleading; we do not record or use the outcome.

Let  $H \in \text{Herm}(\mathcal{H})$  be a single interaction in a many-body system, which we intend to simulate. As before, we will introduce an ancillary qubit  $\mathcal{A} \cong \mathbb{C}^2$ , and evolve under a Hamiltonian  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$ , but now we supplement the natural time evolution with regular projective measurements on the  $\mathcal{A}$  system at time intervals of  $\delta t$ . By the quantum Zeno effect<sup>62</sup>, this forces the  $\mathcal{A}$  system to stay in the  $|0\rangle$  state with high probability, meanwhile simulating the desired interaction on the  $\mathcal{H}$  system.

The following result, Proposition 27, provides a formal construction for the measurement-based gadgets described above — see Appendix B for the proof. Qualitatively, this result tells us that if we evolve  $|\psi\rangle \otimes |0\rangle$  for time  $\delta t$  under the simulator Hamiltonian  $H'$ , and then measure the ancillary qubit, we will obtain a ‘1’ result with probability  $O((\delta t)^3)$  (corresponding to an amplitude of  $O((\delta t)^{3/2})$ ). In the more likely case that we obtain ‘0’, the post-measurement state (on the  $\mathcal{H}$  space) is  $e^{-i\delta t H}|\psi\rangle$ , for some new Hamiltonian  $H$ , up to error  $O((\delta t)^2)$ . By repeating this process  $t/\delta t$  times, we will hence obtain a state  $e^{-itH}|\psi\rangle + O(t(\delta t))$  on the  $\mathcal{H}$  space if ‘0’ is measured in every round of measurement. The probability of a measurement error in this process scales as  $t(\delta t)^2$ , hence can be controlled provided that  $\delta t = O(t^{-1/2})$ , which will always be satisfied if we choose  $\delta t = O(t^{-1})$  in order to control the error on the post-measurement state.

**Proposition 27.** *For a Hilbert space  $\mathcal{H}$  and an ancillary qubit  $\mathcal{A} = \mathbb{C}^2$ , let  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$  be a Hamiltonian given by*

$$H' = H_I \otimes I + H_X \otimes X + H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1| , \quad (54)$$

*for some  $H_I, H_X, H_{|1\rangle\langle 1|} \in \text{Herm}(\mathcal{H})$  depending on a small parameter  $\delta t$  such that  $\|H_I\| = O(1)$ ,  $\|H_X\| = O((\delta t)^{-1/2})$ , and  $\|H_{|1\rangle\langle 1|}\| = O((\delta t)^{-1})$  with  $H_{|1\rangle\langle 1|}^2 = \omega^2 I$ ,  $\omega = \frac{2\pi}{\delta t}$ .*

*Then, for any  $|\psi\rangle \in \mathcal{H}$ ,*

$$e^{-i\delta t H'}(|\psi\rangle \otimes |0\rangle) = (e^{-i\delta t H}|\psi\rangle + O((\delta t)^2)) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle , \quad (55)$$

*where*

$$H = H_I - \omega^{-2} H_X H_{|1\rangle\langle 1|} H_X . \quad (56)$$

This provides a new 3-to-2-local gadget for Pauli strings. For example, we can set  $H_I = -Z_1$ ,  $H_X = \sqrt{\frac{\omega}{2}}(Z_2 + Z_3)$ ,  $H_{|1\rangle\langle 1|} = -\omega Z_1$ ; this yields a 2-local Hamiltonian  $H'$  simulating the 3-local

interaction  $H = Z_1 \otimes Z_2 \otimes Z_3$ . More generally, given three commuting Pauli strings  $A_a, B_b, C_c$ , we can set  $H_I = -A_a$ ,  $H_X = \sqrt{\frac{\omega}{2}}(B_b + C_c)$ ,  $H_{|1\rangle\langle 1|} = -\omega A_a$  to simulate the interaction  $H = A_a \otimes B_b \otimes C_c$ . This procedure may be used to simulate a  $k$ -local Pauli string using a  $(\lceil k/3 \rceil + 1)$ -local Hamiltonian.

Although Proposition 27 shows that evolution and repeated measurements under  $H'$  reproduce the dynamics of  $H$ , it is also important to guarantee that it can be combined with other interactions. Proposition 28 provides the necessary result for this, by verifying that the conclusions of Proposition 27 also hold when an additional term  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$  is added to both the target and simulator Hamiltonian.

**Proposition 28.** *Let  $H_{\text{else}} = \sum_i h_i$  be a  $k$ -local Hamiltonian on  $\mathcal{H} = \otimes_i \mathcal{H}_i$  such that  $\|h_i\| = O(1)$ , and whose interaction graph has a degree bounded by an  $O(1)$  constant.*

*Introduce an ancillary qubit  $\mathcal{A} = \mathbb{C}^2$ , and let  $H' \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$  be a Hamiltonian given by*

$$H' = H_I \otimes \mathbf{I} + H_X \otimes X + H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1| , \quad (57)$$

*for some  $H_I, H_X, H_{|1\rangle\langle 1|} \in \text{Herm}(\mathcal{H})$  depending on a small parameter  $\delta t$  such that  $\|H_I\| = O(1)$ ,  $\|H_X\| = O((\delta t)^{-1/2})$ , and  $\|H_{|1\rangle\langle 1|}\| = O((\delta t)^{-1})$  with  $H_{|1\rangle\langle 1|}^2 = \omega^2 \mathbf{I}$ ,  $\omega = \frac{2\pi}{\delta t}$ . Assume that  $H_I, H_X$ , and  $H_{|1\rangle\langle 1|}$  act on  $O(1)$  sites in  $\mathcal{H}$ .*

*Then, for any  $|\psi\rangle \in \mathcal{H}$ ,*

$$e^{-i\delta t(H' + H_{\text{else}} \otimes \mathbf{I})}(|\psi\rangle \otimes |0\rangle) = (e^{-i\delta t(H + H_{\text{else}})}|\psi\rangle + O((\delta t)^2)) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle , \quad (58)$$

*where*

$$H = H_I - \omega^{-2} H_X H_{|1\rangle\langle 1|} H_X . \quad (59)$$

The significance of Proposition 28 is that the errors do not depend on the size of the system through  $\|H_{\text{else}}\|$ , due to bounds we place on the Trotter error in the expansion  $e^{-i\delta t(H + H_{\text{else}})} \approx e^{-i\delta t H} e^{-i\delta t H_{\text{else}}}$ .

## Discussion

Given the result of Proposition 28, we can now describe how the measurement gadget construction fits into our framework of analogue quantum simulation described in Definition 5.

Given a Hamiltonian  $H = Z_1 \otimes Z_2 \otimes Z_3 + H_{\text{else}}$  on  $n$  qubits  $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ , with  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$  satisfying the requirements of Proposition 28, we fix some  $\delta t > 0$  and define the simulator space  $\mathcal{H}' = \mathcal{H} \otimes \mathcal{A}$ , where  $\mathcal{A} = \mathbb{C}^2$ . Let  $H' \in \text{Herm}(\mathcal{H}')$  be given by

$$H' = -Z_1 \otimes \mathbf{I} + \sqrt{\frac{\omega}{2}}(Z_2 + Z_3) \otimes X - \omega Z_1 \otimes |1\rangle\langle 1| , \quad (60)$$

where  $\omega = \frac{2\pi}{\delta t}$ . Define the state and observable encodings  $\mathcal{E}_{\text{state}}$  and  $\mathcal{E}_{\text{obs}}$  by

$$\mathcal{E}_{\text{state}}(\rho) = \rho \otimes |0\rangle\langle 0| , \quad \mathcal{E}_{\text{obs}}(O) = O \otimes \mathbf{I} , \quad (61)$$

and define channels  $E_{\delta t}, M : D(\mathcal{H}') \rightarrow D(\mathcal{H}')$  by

$$E_{\delta t}(\rho') = e^{-i\delta t(H' + H_{\text{else}} \otimes \mathbf{I})} \rho' e^{i\delta t(H' + H_{\text{else}} \otimes \mathbf{I})} , \quad (62)$$

$$M(\rho') = \text{tr}_{\mathcal{A}}[\rho'(\mathbf{I} \otimes |0\rangle\langle 0|)] \otimes |0\rangle\langle 0| + \text{tr}_{\mathcal{A}}[\rho'(\mathbf{I} \otimes |1\rangle\langle 1|)] \otimes |1\rangle\langle 1| , \quad (63)$$

so that  $E_{\delta t}$  corresponds to evolution under the Hamiltonian  $H' + H_{\text{else}}$  for time  $\delta t$ , and  $M$  corresponds to a measurement of the  $\mathcal{A}$  system. Then, for all  $t$ , define the time evolution channel

$$T_t = (M \circ E_{\delta t}) \circ (M \circ E_{\delta t}) \circ \cdots \circ (M \circ E_{\delta t}) , \quad (64)$$

containing  $\lfloor t/\delta t \rfloor$  copies of  $(M \circ E_{\delta t})$ . This evolution is described by Fig. 4. The content of Proposition 28 tells us that

$$(T_t \circ \mathcal{E}_{\text{state}})(\rho) = (e^{-itH} \rho e^{itH} + O(t\delta t)) \otimes |0\rangle\langle 0| + O(t(\delta t)^2) \otimes |1\rangle\langle 1| , \quad (65)$$

and hence for any observable  $O \in \text{Herm}(\mathcal{H})$  with  $\|O\| = 1$ ,

$$\text{tr}[\mathcal{E}_{\text{obs}}(O)(T_t \circ \mathcal{E}_{\text{state}})(\rho)] = \text{tr}[Oe^{-itH} \rho e^{itH}] + O(t\delta t) . \quad (66)$$

The channels  $T_t$  therefore simulate  $H$  (in the sense of Definition 5) with respect to any states  $\Omega_{\text{state}}$  and normalised observables  $\Omega_{\text{obs}}$ , up to accuracy  $\epsilon > 0$  and maximum time  $t_{\text{max}}$ , provided that one chooses  $\delta t = O(\epsilon t_{\text{max}}^{-1})$ . Therefore we require interaction strengths and measurement frequency which scale as  $J = O(\epsilon^{-1} t_{\text{max}})$  — note that this does not depend on  $n$ , the size of the system.

We can compare these scalings with those obtained if we were to use conventional gadgets. Suppose we have a  $(\eta, \epsilon)$ -gadget in the sense of Definition 15, with  $\eta = O(n^{-1}\epsilon)$  to ensure an absolute error of  $O(\epsilon)$  when combined with a Hamiltonian of order  $n$ , comparable with the above construction. By Theorem 26, this must involve energy scalings of  $J = \Omega(\epsilon^{-1}n)$  (and even without Theorem 26, a low-energy  $(\Delta, \eta, \epsilon)$ -gadget as in Definition 18 would require energies scaling as  $\Omega(n)$  to ensure that unwanted states are sufficiently penalised). In fact, this is likely not the optimal bound; the best known 3-to-2 gadget construction requires energy scales of  $O(\epsilon^{-3} + \eta^{-3})$ , which in this case would require interaction strengths scaling as  $J = O((\epsilon^{-1}n)^3)$ . Even if the system size is restricted via Lieb-Robinson bounds to set  $n = O(\log^d(1/\epsilon) + t_{\text{max}}^d)$  (where  $d$  is the dimension of the system), the measurement-based gadget still provides an improvement.

Despite this advantage, the measurement gadget construction involves repeated instantaneous decoherence of the ancillary qubit at precise time intervals without disturbing the rest of the system, and may still require large (albeit non-scaling) interaction strengths. Moreover, if  $N_{\text{gad}}$  such gadgets were used in parallel, we expect (though do not calculate here) that an additional overhead of at least  $\delta t = O((t_{\text{max}} N_{\text{gad}})^{-1/2})$  would be necessary to control the probability of measuring a 1 at any of the ancillary sites. Nonetheless, the construction provides a marked improvement in scalings over existing gadgets for a single 3-local term in a Hamiltonian, and gives some positive clues as to the ways in which simulators might take advantage of more general possibilities for channels allowed by Definition 5. We leave the detailed study of such gadgets, and their robustness to error for future work. We anticipate that, for a suitable adaptation of Definition 15 for the dissipative case, there may be similar no-go results preventing locality reduction by gadgets independently of the size of the system.

## Acknowledgements

We acknowledge financial support from the Novo Nordisk Foundation (Grant No. NNF20OC0059939 ‘Quantum for Life’), the European Research Council (ERC Grant Agreement No. 818761) and VILLUM FONDEN via the QMATH Centre of Excellence (Grant No. 10059). A.H.W. thanks the VILLUM FONDEN for its support with a Villum Young Investigator Grant (Grant No. 25452). I.D. was supported in part by the AFOSR under grant FA9550-21-1-0392 and a National Science Foundation (NSF) Graduate Research Fellowship under Grant No. DGE 1656518. I.D. thanks everyone at QMATH

for their hospitality during his research visit to KU and especially Prof. Adam Bouland for encouraging and supporting the visit. I.D. gratefully acknowledges Harriet Apel for generously offering insights and guidance during fruitful discussions at the early stages of this work.

## References

- [1] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.
- [2] Dave Wecker, Matthew B. Hastings, Nathan Wiebe, Bryan K. Clark, Chetan Nayak, and Matthias Troyer. Solving strongly correlated electron models on a quantum computer. *Physical Review A*, 92(6):062318, 2015.
- [3] Alberto Baiardi, Matthias Christandl, and Markus Reiher. Quantum computing for molecular biology. *ChemBioChem*, 24(13):e202300120, 2023.
- [4] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.
- [5] Iulia M. Georgescu, Sahel Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153, 2014.
- [6] Dominic W. Berry, Andrew M. Childs, and Robin Kothari. Hamiltonian simulation with nearly optimal dependence on all parameters. In *2015 IEEE 56th Annual Symposium on Foundations of Computer Science*, pages 792–809. IEEE, 2015.
- [7] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019.
- [8] Andrew M. Childs, Dmitri Maslov, Yunseong Nam, Neil J. Ross, and Yuan Su. Toward the first quantum simulation with quantum speedup. *Proceedings of the National Academy of Sciences*, 115(38):9456–9461, 2018.
- [9] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.
- [10] J. Ignacio Cirac and Peter Zoller. Goals and opportunities in quantum simulation. *Nature Physics*, 8(4):264–266, 2012.
- [11] Toby S. Cubitt, Ashley Montanaro, and Stephen Piddock. Universal quantum Hamiltonians. *Proceedings of the National Academy of Sciences*, 115(38):9497–9502, 2018.
- [12] Stephen Piddock and Johannes Bausch. Universal translationally-invariant Hamiltonians. *arXiv preprint arXiv:2001.08050*, 2020.
- [13] Stephen Piddock and Ashley Montanaro. Universal qudit Hamiltonians. *Communications in Mathematical Physics*, 382:721–771, 2021.
- [14] Tamara Kohler, Stephen Piddock, Johannes Bausch, and Toby Cubitt. Translationally invariant universal quantum Hamiltonians in 1D. In *Annales Henri Poincaré*, pages 1–32. Springer, 2020.
- [15] Tamara Kohler, Stephen Piddock, Johannes Bausch, and Toby Cubitt. General conditions for universality of quantum Hamiltonians. *PRX Quantum*, 3(1):010308, 2022.
- [16] Leo Zhou and Dorit Aharonov. Strongly universal Hamiltonian simulators. *arXiv preprint arXiv:2102.02991*, 2021.
- [17] Roberto Oliveira and Barbara M. Terhal. The complexity of quantum spin systems on a two-dimensional square lattice. *Quantum Information and Computation*, 8(10):900–924, 2008.
- [18] Sergey Bravyi and Matthew Hastings. On complexity of the quantum Ising model. *Communications in Mathematical Physics*, 349(1):1–45, 2017.
- [19] Dorit Aharonov and Leo Zhou. Hamiltonian sparsification and gap-simulations. *arXiv preprint arXiv:1804.11084*, 2018.

- [20] Dolev Bluvstein, Harry Levine, Giulia Semeghini, Tout T. Wang, Sepehr Ebadi, Marcin Kalinowski, Alexander Keesling, Nishad Maskara, Hannes Pichler, Markus Greiner, Vladan Vuletić, and Mikhail D. Lukin. A quantum processor based on coherent transport of entangled atom arrays. *Nature*, 604(7906):451–456, 2022.
- [21] Wolfgang Dür, Guifre Vidal, and J. Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62(6):062314, 2000.
- [22] Michael Walter, Brent Doran, David Gross, and Matthias Christandl. Entanglement polytopes: multiparticle entanglement from single-particle information. *Science*, 340(6137):1205–1208, 2013.
- [23] Matthias Christandl, Vladimir Lysikov, Vincent Steffan, Albert H. Werner, and Freek Witteveen. The resource theory of tensor networks. *arXiv preprint arXiv:2307.07394*, 2023.
- [24] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Communications in Mathematical Physics*, 28:251–257, 1972.
- [25] Isaac H. Kim and Brian Swingle. Robust entanglement renormalization on a noisy quantum computer. *arXiv preprint arXiv:1711.07500*, 2017.
- [26] Johannes Borregaard, Matthias Christandl, and Daniel Stilek França. Noise-robust exploration of many-body quantum states on near-term quantum devices. *npj Quantum Information*, 7(1):45, 2021.
- [27] Rahul Trivedi, Adrian Franco Rubio, and J. Ignacio Cirac. Quantum advantage and stability to errors in analogue quantum simulators. *arXiv preprint arXiv:2212.04924*, 2022.
- [28] Vittorio Gorini, Andrzej Kossakowski, and Ennackal Chandy George Sudarshan. Completely positive dynamical semigroups of n-level systems. *Journal of Mathematical Physics*, 17(5):821–825, 1976.
- [29] Goran Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48:119–130, 1976.
- [30] Julia Kempe and Oded Regev. 3-local Hamiltonian is QMA-complete. *Quantum Information and Computation*, 3(3):258–264, 2003.
- [31] Sergey Bravyi, David P. DiVincenzo, Daniel Loss, and Barbara M. Terhal. Quantum simulation of many-body Hamiltonians using perturbation theory with bounded-strength interactions. *Physical Review Letters*, 101(7):070503, 2008.
- [32] Yudong Cao, Ryan Babbush, Jacob Biamonte, and Sabre Kais. Hamiltonian gadgets with reduced resource requirements. *Physical Review A*, 91(1):012315, 2015.
- [33] Yudong Cao and Daniel Nagaj. Perturbative gadgets without strong interactions. *Quantum Information and Computation*, 15(13-14):1197–1222, 2015.
- [34] Simon Cichy, Paul K. Faehrmann, Sumeet Khatri, and Jens Eisert. A perturbative gadget for delaying the onset of barren plateaus in variational quantum algorithms. *arXiv preprint arXiv:2210.03099*, 2022.
- [35] Philippe Lewalle, Leigh S. Martin, Emmanuel Flurin, Song Zhang, Eliya Blumenthal, Shay Hacoheh-Gourgy, Daniel Burgarth, and K. Birgitta Whaley. A multi-qubit quantum gate using the Zeno effect. *Quantum*, 7:1100, 2023.
- [36] Carter Ball and Thomas D. Cohen. Zeno effect suppression of gauge drift in quantum simulations. *arXiv preprint arXiv:2405.09462*, 2024.
- [37] Eliya Blumenthal, Chen Mor, Asaf A. Diringer, Leigh S. Martin, Philippe Lewalle, Daniel Burgarth, K. Birgitta Whaley, and Shay Hacoheh-Gourgy. Demonstration of universal control between non-interacting qubits using the Quantum Zeno effect. *npj Quantum Information*, 8(1):88, 2022.
- [38] David P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000.
- [39] A. Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997.
- [40] Henrik Wilming and Albert H. Werner. Lieb-Robinson bounds imply locality of interactions. *Physical Review B*, 105(12):125101, 2022.

- [41] Alexei Yu Kitaev, Alexander Shen, and Mikhail N. Vyalyi. *Classical and quantum computation*. American Mathematical Society, 2002.
- [42] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local Hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.
- [43] Sean Hallgren, Daniel Nagaj, and Sandeep Narayanaswami. The local Hamiltonian problem on a line with eight states is QMA-complete. *Quantum Information and Computation*, 13(9-10):721–750, 2013.
- [44] Toby Cubitt and Ashley Montanaro. Complexity classification of local Hamiltonian problems. *SIAM Journal on Computing*, 45(2):268–316, 2016.
- [45] Daniel Gottesman and Sandy Irani. The quantum and classical complexity of translationally invariant tiling and Hamiltonian problems. In *2009 50th Annual IEEE Symposium on Foundations of Computer Science*, pages 95–104. IEEE, 2009.
- [46] Harriet Apel and Toby Cubitt. A mathematical framework for quantum Hamiltonian simulation and duality. *arXiv preprint arXiv:2208.11941*, 2022.
- [47] Marco Cerezo, Andrew Arrasmith, Ryan Babbush, Simon C. Benjamin, Suguru Endo, Keisuke Fujii, Jarrod R. McClean, Kosuke Mitarai, Xiao Yuan, Lukasz Cincio, et al. Variational quantum algorithms. *Nature Reviews Physics*, 3(9):625–644, 2021.
- [48] Jarrod R. McClean, Jonathan Romero, Ryan Babbush, and Alán Aspuru-Guzik. The theory of variational hybrid quantum-classical algorithms. *New Journal of Physics*, 18(2):023023, 2016.
- [49] Frank Verstraete, Michael M. Wolf, and J. Ignacio Cirac. Quantum computation and quantum-state engineering driven by dissipation. *Nature Physics*, 5(9):633–636, 2009.
- [50] Andrew M. Childs, David Gosset, and Zak Webb. Universal computation by multiparticle quantum walk. *Science*, 339(6121):791–794, 2013.
- [51] Thomas C. Bohdanowicz and Fernando G.S.L. Brandão. Universal Hamiltonians for exponentially long simulation. *arXiv preprint arXiv:1710.02625*, 2017.
- [52] Terry Farrelly. A review of quantum cellular automata. *Quantum*, 4:368, 2020.
- [53] Wolfgang Lechner, Philipp Hauke, and Peter Zoller. A quantum annealing architecture with all-to-all connectivity from local interactions. *Science advances*, 1(9):e1500838, 2015.
- [54] Minh-Thi Nguyen, Jin-Guo Liu, Jonathan Wurtz, Mikhail D Lukin, Sheng-Tao Wang, and Hannes Pichler. Quantum optimization with arbitrary connectivity using Rydberg atom arrays. *PRX Quantum*, 4(1):010316, 2023.
- [55] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.
- [56] Jacob D. Biamonte and Peter J. Love. Realizable Hamiltonians for universal adiabatic quantum computers. *Physical Review A*, 78(1):012352, 2008.
- [57] Norbert Schuch and Frank Verstraete. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nature physics*, 5(10):732–735, 2009.
- [58] Stephen Piddock and Ashley Montanaro. The complexity of antiferromagnetic interactions and 2D lattices. *Quantum Information and Computation*, 17(7-8):636–672, 2017.
- [59] Sergey Bravyi, David P. DiVincenzo, and Daniel Loss. Schrieffer–Wolff transformation for quantum many-body systems. *Annals of Physics*, 326(10):2793–2826, 2011.
- [60] Stephen P. Jordan and Edward Farhi. Perturbative gadgets at arbitrary orders. *Physical Review A*, 77(6):062329, 2008.
- [61] Johannes Bausch. Perturbation gadgets: Arbitrary energy scales from a single strong interaction. *Annales Henri Poincaré*, 21(1):81–114, 2020.
- [62] Baidyanath Misra and E. C. George Sudarshan. The Zeno’s paradox in quantum theory. *Journal of Mathematical Physics*, 18(4):756–763, 1977.



- [63] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science and Business Media, 2013.
- [64] Chandler Davis and William M. Kahan. Some new bounds on perturbation of subspaces. *Bulletin of the American Mathematical Society*, 75(4):863–868, 1969.
- [65] Chandler Davis and William Morton Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.
- [66] Andrew M. Childs, Yuan Su, Minh C. Tran, Nathan Wiebe, and Shuchen Zhu. Theory of Trotter error with commutator scaling. *Physical Review X*, 11(1):011020, 2021.

## A Local encodings

In this section we prove the simple results concerning local encodings.

*Proof of Proposition 7.* We prove this for a single Kraus operator  $X'_k$  acting on  $O(1)$  sites in  $\mathcal{H}'$ , and the result follows by linearity. Note that

$$\begin{aligned} X'_k \mathcal{E}_{\text{state}}(\rho)(X'_k)^\dagger &= X'_k \text{tr}_E[U(\rho \otimes |0\rangle\langle 0|_F)U^\dagger](X'_k)^\dagger \\ &= \text{tr}_E[UX_k(\rho \otimes |0\rangle\langle 0|_F)X_k^\dagger U^\dagger], \end{aligned} \quad (67)$$

where  $X_k = U^\dagger(X'_k \otimes \mathbb{I}_E)U$ , which acts on  $O(1)$  sites in  $\mathcal{H} \otimes F$  by the causality assumptions on  $U$ . ■

*Proof of Proposition 9.* To see this, we use the definition of local state encodings and write

$$\mathcal{E}_{\text{obs}}^*(\rho') = \text{tr}_G[W(\rho' \otimes |0\rangle\langle 0|_E)W^\dagger], \quad (68)$$

where  $W \in \text{U}(\mathcal{H}' \otimes E, \mathcal{H} \otimes G)$  is a constant-depth quantum circuit. Then the measurement expectation value is

$$\begin{aligned} \text{tr}[\mathcal{E}_{\text{obs}}(O)\rho'] &= \text{tr}[O\mathcal{E}_{\text{obs}}^*(\rho')] \\ &= \text{tr}[(O \otimes \mathbb{I}_G)W(\rho' \otimes |0\rangle\langle 0|_E)W^\dagger] \\ &= \text{tr}[(\mathbb{I}_{\mathcal{H}'} \otimes \langle 0|_E)W^\dagger(O \otimes \mathbb{I}_G)W(\mathbb{I}_{\mathcal{H}'} \otimes |0\rangle_E)\rho']. \end{aligned} \quad (69)$$

Assuming  $O$  is local, then  $W^\dagger(O \otimes \mathbb{I}_G)W$  acts only on a constant-sized subsystem of  $\mathcal{H}'$ . In particular, we can write  $\mathcal{H}' = A \otimes A^c$  where  $A$  consists of  $O(1)$  sites, and then

$$W^\dagger(O \otimes \mathbb{I}_G)W = O' \otimes \mathbb{I}_{A^c}, \quad (70)$$

for some  $O'$  acting on  $A \otimes E$ . Then

$$\text{tr}[\mathcal{E}_{\text{obs}}(O)\rho'] = \text{tr}[(\mathbb{I}_A \otimes \langle 0|_E)O'(\mathbb{I}_A \otimes |0\rangle_E)\rho'_A], \quad (71)$$

which can be estimated via a POVM on  $A$ . ■

## B Qutrit-to-qubit energy scaling

The idea for the proof of Proposition 10 is simple: by encoding a qutrit into a set of qubits, we must end up with an “unused” state in the qubit system, since the encoding cannot be surjective by dimension counting. Since the  $(\Delta, \eta, \epsilon)$  simulation requires all simulated states to lie in the low-energy subspace of the simulator, this implies that the unused qubit states must lie in the high-energy (above  $\Delta$ ) subspace.

In the proof below, we start with the encoded ground state  $\rho_0$ , and construct a state  $\rho_1$  which differs only from  $\rho_0$  only in one set of qubits in which it is in such an “unused” state. The similarity of the states and their differences in energies lead to the requirement for strong interactions.

In this proof, and subsequent sections, we make frequent use of the following standard result from matrix analysis<sup>63</sup>.

**Lemma 29** (Weyl’s Perturbation Theorem). *Let  $A, B \in \text{Herm}(\mathcal{H})$  be Hermitian matrices, with spectra  $\lambda_0 \leq \lambda_1 \leq \dots$  and  $\mu_0 \leq \mu_1 \leq \dots$  respectively. Then*

$$\max_j |\lambda_j - \mu_j| \leq \|A - B\| . \quad (72)$$

*Proof of Proposition 10.* Write  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$ , where  $\mathcal{H}_i = \mathbb{C}^3$  is a single qutrit site. By the definition of local simulation given by Cubitt et al.<sup>11</sup>, we have two encodings  $\mathcal{E}_{\text{obs}}$  and  $\tilde{\mathcal{E}}_{\text{obs}}$  of the form (using that  $H$  is real to set  $Q = 0$  without loss of generality)

$$\mathcal{E}_{\text{obs}}(M) = V(M \otimes P)V^\dagger , \quad \tilde{\mathcal{E}}_{\text{obs}}(M) = \tilde{V}(M \otimes P)\tilde{V}^\dagger , \quad (73)$$

where  $P$  is a projector on the ancillary space  $\mathcal{A}$ , and  $V, \tilde{V}$  are both isometries  $\mathcal{H} \otimes \mathcal{A} \rightarrow \mathcal{H}'$ . These encodings satisfy the properties:

- $\mathcal{E}_{\text{obs}}$  is a local encoding, in the sense that  $\mathcal{A} = \otimes_{i=1}^n \mathcal{A}_i$  and  $V = \otimes_{i=1}^n V_i$  where  $V_i : \mathcal{H}_i \otimes \mathcal{A}_i \rightarrow \mathcal{H}'_i$ . Here we write  $\mathcal{H}'_i \cong (\mathbb{C}^2)^{\otimes m_i}$  for the set of  $m_i$  qubits into which qutrit  $i$  is encoded. Note  $\sum_i m_i = m$ .

- $\tilde{\mathcal{E}}_{\text{obs}}$  satisfies

$$\tilde{\mathcal{E}}_{\text{obs}}(\mathbf{I}) = \tilde{V}(\mathbf{I} \otimes P)\tilde{V}^\dagger = P_{\leq \Delta(H'_n)} , \quad (74)$$

where  $P_{\leq \Delta(H'_n)}$  is the low-energy (below  $\Delta$ ) projector for  $H'_n$ , and

$$\|P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} - \tilde{\mathcal{E}}_{\text{obs}}(H_n)\| \leq \epsilon . \quad (75)$$

- $\mathcal{E}_{\text{obs}}$  and  $\tilde{\mathcal{E}}_{\text{obs}}$  are close, in the sense that

$$\|V - \tilde{V}\| \leq \eta . \quad (76)$$

Now we define a state  $\tau \in \text{span}(P)$  and define a state encoding (in the sense of Cubitt et al.<sup>11</sup>)

$$\tilde{\mathcal{E}}_{\text{state}}(\rho) = \tilde{V}(\rho \otimes \tau)\tilde{V}^\dagger . \quad (77)$$

Let  $\rho_0 = \tilde{\mathcal{E}}_{\text{state}}(|\downarrow\rangle\langle\downarrow|^{\otimes n})$  be the encoded ground state of  $H_n$ , which by definition satisfies

$$P_{\leq \Delta(H'_n)} \rho_0 = \rho_0 , \quad \tilde{\mathcal{E}}_{\text{obs}}(H_n) \rho_0 = 0 . \quad (78)$$

Hence we can bound the energy of  $\rho_0$  under  $H'_n$  by

$$\begin{aligned}
\text{tr}[H'_n \rho_0] &= \text{tr}[P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} \rho_0] \\
&= \text{tr}[(P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} - \tilde{\mathcal{E}}_{\text{obs}}(H_n)) \rho_0] \\
&\leq \|P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} - \tilde{\mathcal{E}}_{\text{obs}}(H_n)\| \\
&\leq \epsilon .
\end{aligned} \tag{79}$$

Now without loss of generality we assume that  $m_1 = \min_i m_i$ . Notice that  $V_1 : \mathcal{H}_1 \otimes \mathcal{A}_1 \rightarrow \mathcal{H}'_1$  cannot be surjective, since

$$\dim(\mathcal{H}_1 \otimes \mathcal{A}_1) = 3 \dim \mathcal{A}_1 \neq 2^{m_1} . \tag{80}$$

We can therefore choose some pure state  $\psi = |\psi\rangle\langle\psi|$  in  $\mathcal{H}'_1$  which is orthogonal to the image of  $V_1$ , and define

$$\rho_1 = \psi \otimes \text{tr}_1[\rho_0] \in \text{Lin}(\mathcal{H}') , \tag{81}$$

Where  $\text{tr}_1$  denotes the partial trace over the  $\mathcal{H}'_1$  system. This satisfies  $V^\dagger \rho_1 = \rho_1 V = 0$ , so we have

$$\begin{aligned}
\text{tr}[P_{\leq \Delta(H'_n)} \rho_1] &= \text{tr}[(\mathbf{I} \otimes P) \tilde{V}^\dagger \rho_1 \tilde{V}] \\
&\leq \text{tr}[\tilde{V}^\dagger \rho_1 \tilde{V}] \\
&= \text{tr}[(\tilde{V} - V)^\dagger \rho_1 (\tilde{V} - V)] \\
&\leq \|(\tilde{V} - V)(\tilde{V} - V)^\dagger\| \\
&\leq \eta^2 ,
\end{aligned} \tag{82}$$

from which we deduce that

$$\text{tr}[H'_n \rho_1] \geq \Delta \text{tr}[(\mathbf{I} - P_{\leq \Delta(H'_n)}) \rho_1] - \epsilon \text{tr}[P_{\leq \Delta(H'_n)} \rho_1] \geq \Delta(1 - \eta^2) - \epsilon \eta^2 , \tag{83}$$

using that the smallest eigenvalue of  $H'_n$  is at least  $-\epsilon$ , by (75) and Lemma 29. Therefore, using (79),

$$\text{tr}[H'_n(\rho_1 - \rho_0)] \geq \Delta(1 - \eta^2) - \epsilon(1 + \eta^2) . \tag{84}$$

On the other hand, by expanding  $H'_n$  we can write

$$\text{tr}[H'_n(\rho_1 - \rho_0)] = \sum_{j=1}^K \text{tr}[h'_j(\rho_1 - \rho_0)] . \tag{85}$$

Notice that if  $h'_j$  acts trivially on  $\mathcal{H}'_1$ , that is  $h'_j = \mathbf{I}_1 \otimes \tilde{h}_j$ , then

$$\begin{aligned}
\text{tr}[h'_j(\rho_1 - \rho_0)] &= \text{tr}[(\mathbf{I}_1 \otimes \tilde{h}_j)(\psi \otimes \text{tr}_1 \rho_0 - \rho_0)] \\
&= \text{tr}_1[\psi] \text{tr}_{2,3,\dots}[\tilde{h}_j \text{tr}_1[\rho_0]] - \text{tr}[(\mathbf{I} \otimes \tilde{h}_j) \rho_0] \\
&= 0 .
\end{aligned} \tag{86}$$

Hence the only non-zero contributions to (85) come from  $j$  in the set

$$I_1 = \{1 \leq j \leq K \mid h'_j \text{ acts non-trivially on } \mathcal{H}'_1\} . \tag{87}$$

So (85) can be bounded by

$$\mathrm{tr}[H'_n(\rho_1 - \rho_0)] = \sum_{j \in I_1} \mathrm{tr}[h'_j(\rho_1 - \rho_0)] \leq 2|I_1| \max_{j \in I_1} \|h'_j\|, \quad (88)$$

using the Hölder inequality for Schatten  $p$ -norms.

Now notice that, since the largest eigenvalue of  $H_n$  is  $n$ , and the encoding  $\tilde{\mathcal{E}}_{\mathrm{obs}}$  preserves spectra, we have

$$\|\tilde{\mathcal{E}}_{\mathrm{obs}}(H_n)\| = \|H_n\| = n, \quad (89)$$

so by (75) and Lemma 29

$$\|P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)}\| \geq n - \epsilon. \quad (90)$$

Hence, by the definition of  $P_{\leq \Delta(H'_n)}$ , we must have  $\Delta > n - \epsilon$ . Combining this fact with (84) and (88), we deduce that

$$\max_{j \in I_1} \|h'_j\| > \frac{1}{2|I_1|} ((n - \epsilon)(1 - \eta^2) - \epsilon(1 + \eta^2)). \quad (91)$$

Finally, note that  $m_1 \leq m/n = O(n^\alpha)$  and  $|I_1| \leq dm_1$ , so for large  $n$  we have the desired scaling

$$\max_{j \in I_1} \|h'_j\| \geq \Omega(n^{1-\alpha}(1 - \eta^2)). \quad (92)$$

■

*Proof of Lemma 11.* A ball of radius  $R$  contains  $O(R^D)$  points in  $\mathbb{Z}^D$ , so at most  $O(R^D/n^\alpha) = O(n^{1-\alpha})$  of the  $X_i$  can be fully contained within.

If an  $X_i$  is partially contained within the ball, then it must intersect with its boundary. There are only  $O(R^{D-1})$  points on the boundary, so this can be the case for  $O(R^{D-1}) = O(n^{1-1/D})$  of the  $X_i$ .

Hence the total number of  $X_i$  that can be fully or partially contained within a ball of radius  $R$  is upper bounded by

$$|B_R(y)| = O(n^{1-\alpha} + n^{1-1/D}) = O(n^{1-\min\{\alpha, 1/D\}}). \quad (93)$$

■

## C Gadget characterisation

In this section, we give the proofs of Theorems 16 and 17, showing the equivalence of our notions of gadgets. The former is quite simple, but the latter requires several preparatory lemmas. In particular, we will make heavy use of the direct rotation — for a detailed introduction see the review of Bravyi et al.<sup>59</sup>. We summarise the basic definitions and properties here without proof.

### C.1 The direct rotation

Consider two states  $|\psi\rangle, |\phi\rangle$  lying in some Hilbert space  $\mathcal{H} \cong \mathbb{C}^N$ . There are many unitary matrices  $U \in \mathrm{U}(\mathcal{H})$  which rotate between these states (that is,  $U|\psi\rangle = |\phi\rangle$ ), but a particularly natural choice is the unitary  $U_{\psi \rightarrow \phi}$  which rotates only within the subspace spanned by  $|\psi\rangle$  and  $|\phi\rangle$ . Defining the

reflections  $R_\psi = \mathbb{I} - 2|\psi\rangle\langle\psi|$  and  $R_\phi = \mathbb{I} - 2|\phi\rangle\langle\phi|$ , we can write

$$U_{\psi \rightarrow \phi} = \sqrt{R_\phi R_\psi}, \quad (94)$$

assuming it is well-defined. This is the direct rotation from  $|\psi\rangle$  to  $|\phi\rangle$ .

This construction can be generalised to rotations between subspaces:

**Definition 30** (Direct rotation). Let  $\mathcal{P}$  and  $\mathcal{Q}$  be linear subspaces of equal dimension corresponding to orthogonal projectors  $P$  and  $Q$  respectively. Define

$$R_{\mathcal{P}} = \mathbb{I} - 2P, \quad R_{\mathcal{Q}} = \mathbb{I} - 2Q, \quad (95)$$

then direct rotation between  $\mathcal{P}$  and  $\mathcal{Q}$  is

$$U_{\mathcal{P} \rightarrow \mathcal{Q}} = \sqrt{R_{\mathcal{Q}} R_{\mathcal{P}}}, \quad (96)$$

where the square root is taken with a branch cut along the negative axis and such that  $\sqrt{1} = 1$ . This is well-defined whenever  $\|P - Q\| < 1$ .

Then  $U_{\mathcal{P} \rightarrow \mathcal{Q}}$  satisfies

$$U_{\mathcal{P} \rightarrow \mathcal{Q}} P U_{\mathcal{P} \rightarrow \mathcal{Q}}^\dagger = Q. \quad (97)$$

Moreover, as described by Bravyi et al.<sup>59</sup>, the direct rotation may be written in terms of its generator: an anti-Hermitian operator  $S = -S^\dagger$  which can be chosen so that  $U_{\mathcal{P} \rightarrow \mathcal{Q}} = e^S$ , with  $\|S\| < \pi/2$  and which is off-diagonal with respect to both  $P$  and  $Q$ :

$$PSP = (\mathbb{I} - P)S(\mathbb{I} - P) = QSQ = (\mathbb{I} - Q)S(\mathbb{I} - Q) = 0. \quad (98)$$

Notice that, writing  $S = i \operatorname{diag}(\theta_1, \theta_2, \dots, \theta_n)$  for  $\theta_j \in (-\pi/2, \pi/2)$ , we have

$$\|U_{\mathcal{P} \rightarrow \mathcal{Q}} - \mathbb{I}\| = \max_j |2 \sin(\theta_j/2)|, \quad \|S\| = \max_j |\theta_j|, \quad (99)$$

and hence

$$\|S\| \leq \frac{\pi}{2\sqrt{2}} \|U_{\mathcal{P} \rightarrow \mathcal{Q}} - \mathbb{I}\|. \quad (100)$$

## C.2 The gadget property from gadgets

*Proof of Theorem 16.* This follows directly from the definition, since for any  $H_{\text{else}} \in \operatorname{Herm}(\mathcal{H})$ , we have

$$\begin{aligned} & \|P'(H' + H_{\text{else}} \otimes \mathbb{I})P' - U((H + H_{\text{else}}) \otimes P)U^\dagger\| \\ & \leq \|P'H'P' - U(H \otimes P)U^\dagger\| + \|P'(H_{\text{else}} \otimes \mathbb{I})P' - U(H_{\text{else}} \otimes P)U^\dagger\|. \end{aligned} \quad (101)$$

The first term is bounded by  $\epsilon$  by definition, and the second term can be bounded using

$$\begin{aligned} & \|P'(H_{\text{else}} \otimes \mathbb{I})P' - U(H_{\text{else}} \otimes P)U^\dagger\| \\ & = \|(\mathbb{I} \otimes P)U^\dagger(H_{\text{else}} \otimes \mathbb{I})U(\mathbb{I} \otimes P) - H_{\text{else}} \otimes P\| \\ & \leq 2\eta \|H_{\text{else}}\|. \end{aligned} \quad (102)$$

Hence the gadget property is satisfied, putting  $\tilde{P} = P$  and  $\tilde{U}_{H_{\text{else}}} := U$  for all  $H_{\text{else}}$ .  $\blacksquare$

### C.3 Gadgets from the gadget property

The proof of Theorem 17 requires Lemma 31, a basic linear algebra fact which we prove here for convenience. Two projectors  $P$  and  $Q$  commute if and only if they are simultaneously diagonalisable, in which case  $PQP$  is also a projector. This lemma says that this is also true in the approximate setting:  $[P, Q]$  is small if and only if  $PQP$  is close to some projector  $\tilde{P}$ .

**Lemma 31.** *Let  $P, Q \in \text{Proj}(\mathcal{H})$  be projectors, and define*

$$f(P, Q) := \min_{\tilde{P} \in \text{Proj}(\mathcal{H})} \|PQP - \tilde{P}\|. \quad (103)$$

Then

$$\|[P, Q]\| = \sqrt{f(P, Q) - f(P, Q)^2}. \quad (104)$$

*Proof of Lemma 31.* Write  $P$  and  $Q$  in the block-diagonal basis of  $P$ , so that

$$P = \begin{pmatrix} \mathbf{I} & 0 \\ 0 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix}, \quad (105)$$

for some matrices  $A, B, C$ . The requirement  $Q \in \text{Proj}(\mathcal{H})$  implies that  $BB^\dagger = A(\mathbf{I} - A)$ . We have

$$PQP = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix}. \quad (106)$$

Let  $\{\lambda_j\}_j$  be the eigenvalues of  $A$ ; notice that these satisfy  $0 \leq \lambda_j \leq 1$ , since  $0 \leq PQP \leq Q$ . Then  $f(P, Q)$  is given by

$$f(P, Q) = \max_j (\min\{|\lambda_j|, |1 - \lambda_j|\}). \quad (107)$$

To see why (107) holds, note that the upper bound on  $f$  follows by constructing  $\tilde{P}$  to have the same eigenvectors as  $PQP$ , but with each eigenvalue replaced by either 0 or 1 depending on which is closer. The lower bound follows from Lemma 29.

Now we can compute

$$-[P, Q]^2 = \begin{pmatrix} BB^\dagger & 0 \\ 0 & B^\dagger B \end{pmatrix}, \quad (108)$$

hence

$$\|[P, Q]\|^2 = \|[P, Q]^2\| = \|BB^\dagger\| = \|A(\mathbf{I} - A)\| = \max_j |\lambda_j| |1 - \lambda_j|. \quad (109)$$

Note that the maximising  $j$  in (107) and (109) must be the same (the functions  $\min\{|\lambda|, |1 - \lambda|\}$  and  $|\lambda||1 - \lambda|$  are both maximised by the  $\lambda_j$  closest to  $1/2$ ), hence we can deduce

$$\|[P, Q]\|^2 = f(P, Q)(1 - f(P, Q)), \quad (110)$$

which gives the result. ■

In order to obtain the correct unitary  $U$  in the gadget definition, the proof of Theorem 17 requires constructing rotations between eigenspaces of different operators. The Davis-Kahan  $\sin \theta$  theorem below provides a bound on the size of these rotations. This is also used in the proof of Proposition 24.

**Lemma 32** (Davis-Kahan  $\sin \theta$  theorem<sup>64</sup>). *Let  $A, B \in \text{Herm}(\mathcal{H})$ , and take  $P_A, P_B \in \text{Proj}(\mathcal{H})$  projectors of equal rank which block-diagonalise  $A$  and  $B$  respectively, so that*

$$A = P_A A P_A + P_A^\perp A P_A^\perp, \quad B = P_B B P_B + P_B^\perp B P_B^\perp. \quad (111)$$

Assume  $\alpha, \beta \in \mathbb{R}$  and  $\lambda_{\text{gap}}$  are such that

$$\text{spec}(A|_{P_A \mathcal{H}}) \subset [\alpha, \beta], \quad \text{spec}(B|_{P_B^\perp \mathcal{H}}) \subset \mathbb{R} \setminus (\alpha - \lambda_{\text{gap}}, \beta + \lambda_{\text{gap}}). \quad (112)$$

Then the direct rotation  $U \in \text{U}(\mathcal{H})$  from  $P_A$  to  $P_B$  satisfies

$$\|U - \mathbf{I}\| \leq \frac{\sqrt{2}}{\lambda_{\text{gap}}} \|(B - A)P_A\|. \quad (113)$$

*Proof of Lemma 32.* The statement of Davis et al.<sup>64</sup> is phrased in terms of a matrix  $\Theta_0 = \text{diag}(\theta_1, \theta_2, \dots, \theta_n)$ , where the eigenvalues (possibly excluding some 1's) of  $U$  are given by  $e^{i\theta_j}$  and  $\pi/2 \geq \theta_1 \geq \theta_2 \geq \dots \geq \theta_n$ . Specifically, the authors give the following result:

$$\lambda_{\text{gap}} \|\sin \Theta_0\| \leq \|(B - A)P_A\|. \quad (114)$$

To recover our restatement of the theorem, we use the identity  $|1 - e^{i\theta}| = |2 \sin(\theta/2)|$  to deduce that

$$\|U - \mathbf{I}\| = |2 \sin(\theta_1/2)| \leq \sqrt{2} |\sin \theta_1| = \sqrt{2} \|\sin \Theta_0\|. \quad (115)$$

■

The following lemmas from Bravyi et al.<sup>31</sup> are used extensively in the rest of the gadget proofs. They provide bounds on a series expansion of  $e^S H e^{-S}$ , for  $S$  small and anti-Hermitian, in particular showing that

$$e^S H e^{-S} = H + [S, H] + \frac{1}{2!} [S, [S, H]] + \frac{1}{3!} [S, [S, [S, H]]] + \dots \quad (116)$$

**Lemma 33** (Bravyi et al., Lemma 1<sup>31</sup>). *Let  $S$  be an anti-Hermitian operator. Define a superoperator  $\text{ad}_S$  such that  $\text{ad}_S(X) = [S, X]$ , and let  $\text{ad}_S^k$  be the  $k$ -fold composition of  $\text{ad}_S$ , with  $\text{ad}_S^0(X) = X$ . For any operator  $H$  define  $r_0(H) = \|e^S H e^{-S}\| = \|H\|$ ,  $r_1(H) = \|e^S H e^{-S} - H\|$ , and*

$$r_k(H) = \|e^S H e^{-S} - \sum_{p=0}^{k-1} \frac{1}{p!} \text{ad}_S^p(H)\|, \quad k \geq 2. \quad (117)$$

Then for all  $k \geq 0$  one has

$$r_k(H) \leq \frac{1}{k!} \|\text{ad}_S^k(H)\|. \quad (118)$$

**Lemma 34** (Bravyi et al., Lemma 2<sup>31</sup>). *Let  $S = \sum_i S_i$  and  $H = \sum_j H_j$  be any  $O(1)$ -local operators acting on  $n$  qubits with interaction strengths  $J_S$  and  $J_H$  respectively (i.e.  $\|S_i\| \leq J_S$  and  $\|H_j\| \leq J_H$  for all  $i$  and  $j$ ). Let each qubit be acted on non-trivially by  $O(1)$  terms in both  $S$  and  $H$ . Then, for any  $k = O(1)$ ,*

$$\|\text{ad}_S^k(H)\| = O(n J_S^k J_H). \quad (119)$$

These lemmas provide us with the necessary tools to prove Theorem 17.

*Proof of Theorem 17.* The idea of the proof is as follows.

By the gadget property, we have

$$\|P'(H' + H_{\text{else}} \otimes \mathbf{I})P' - \tilde{U}_{H_{\text{else}}}((H + H_{\text{else}}) \otimes \tilde{P})\tilde{U}_{H_{\text{else}}}^\dagger\| \leq \epsilon + \zeta \|H_{\text{else}}\|, \quad (120)$$

for any  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$ .

1. First we consider (120) the case where  $H_{\text{else}}$  dominates the expression, and argue that that  $P'$  is “almost” a projector  $\mathbf{I} \otimes P$  on  $\mathcal{A}$ .
2. Next, by setting  $H_{\text{else}} = 0$  in (120), we observe that  $P'H'P'$  has approximately the same spectrum as  $H \otimes P$ .
3. By setting  $H_{\text{else}} = -H$  in (120), we argue that  $P'H'P' \approx H \otimes P$ .
4. Using steps 2 and 3, and Lemma 32, we construct a rotation  $U$  such that  $\|P'H'P' - U(H \otimes P)U^\dagger\| \leq \epsilon$ , by inductively rotating each eigenspace.

Here we start step 1. Assume that  $\|H_{\text{else}}\| = 1$ , and let  $\lambda > 0$ . Then putting  $H_{\text{else}} \mapsto \lambda H_{\text{else}}$  in (120) yields

$$\|P'(H_{\text{else}} \otimes \mathbf{I})P' - \tilde{U}_{\lambda H_{\text{else}}}(H_{\text{else}} \otimes \tilde{P})\tilde{U}_{\lambda H_{\text{else}}}^\dagger\| \leq \zeta + O(\lambda^{-1}), \quad (121)$$

for large  $\lambda$ , which in particular, by Lemma 29, implies that  $P'(H_{\text{else}} \otimes \mathbf{I})P'$  has the same spectrum as  $H_{\text{else}} \otimes \tilde{P}$ , up to error  $\zeta$ . (That is, the  $k$ th smallest eigenvalue, counted with multiplicity, of  $P'(H_{\text{else}} \otimes \mathbf{I})P'$ , differs from that of  $H_{\text{else}} \otimes \tilde{P}$  by an absolute error of at most  $\zeta$ .)

In particular, if  $H_{\text{else}} = Q \in \text{Proj}(\mathcal{H})$  is a projection, then so is  $P'(Q \otimes \mathbf{I})P'$  (up to spectral error  $\zeta$ ). Hence by Lemma 31, we have

$$\|[P', Q \otimes \mathbf{I}]\| \leq \sqrt{\zeta}. \quad (122)$$

Without loss of generality we may assume that  $\dim \mathcal{H} = K = O(1)$ , by disregarding all the systems on which  $H$ ,  $H'$ , and  $P'$  do not act, using the assumptions of the theorem. So by writing  $H_{\text{else}}$  as a linear combination of at most  $K$  projections, we have

$$\|[P', H_{\text{else}} \otimes \mathbf{I}]\| \leq \pi K \sqrt{\zeta}, \quad (123)$$

for all  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$  with  $\|H_{\text{else}}\| \leq \pi$ .

Therefore, for any  $V \in U(\mathcal{H})$ , we can write  $V = e^{iH_{\text{else}}}$  for some  $H_{\text{else}}$  as above, and then by Lemma 33,

$$\|(V \otimes \mathbf{I})P'(V^\dagger \otimes \mathbf{I}) - P'\| \leq \pi K \sqrt{\zeta}. \quad (124)$$

Integrating over all  $V \in U(\mathcal{H})$  using the Haar measure (normalised with  $\int dV = 1$ ) yields

$$\left\| \int dV (V \otimes \mathbf{I})P'(V^\dagger \otimes \mathbf{I}) - P' \right\| \leq \pi K \sqrt{\zeta}, \quad (125)$$

but

$$\int dV (V \otimes \mathbf{I})P'(V^\dagger \otimes \mathbf{I}) = K^{-1} \mathbf{I}_{\mathcal{H}} \otimes \text{tr}_{\mathcal{H}}[P'], \quad (126)$$

hence

$$\|P' - \mathbf{I}_{\mathcal{H}} \otimes K^{-1} \text{tr}_{\mathcal{H}}[P']\| \leq \pi K \sqrt{\zeta}. \quad (127)$$



In particular, by Lemma 29 this implies that  $K^{-1} \text{tr}_{\mathcal{H}}[P']$  has spectrum in  $[-\pi K\sqrt{\zeta}, \pi K\sqrt{\zeta}] \cup [1 - \pi K\sqrt{\zeta}, 1 + \pi K\sqrt{\zeta}]$  (where for sufficiently small  $\zeta$  there will be a gap). We can therefore construct a projector  $P \in \text{Proj}(\mathcal{A})$  by rounding the eigenvalues of  $K^{-1} \text{tr}_{\mathcal{H}}[P']$  to the nearest integer, which satisfies

$$\|P' - \text{I} \otimes P\| \leq 2\pi K\sqrt{\zeta}. \quad (128)$$

Now we can apply Lemma 32, using  $A = P_A = P'$ ,  $B = P_B = \text{I} \otimes P$ , and  $\lambda_{\text{gap}} = 1$ . Then for  $\zeta$  small enough, the direct rotation  $W$  from  $\text{I} \otimes P$  to  $P'$  satisfies

$$P' = W(\text{I} \otimes P)W^\dagger, \quad \|W - \text{I}\| \leq 2\sqrt{2}\pi K\sqrt{\zeta}, \quad (129)$$

which completes step 1. Without loss of generality we can now adjust the  $\tilde{U}_{H_{\text{else}}}$  so that  $\tilde{P} = P$ , since (121) holds for all  $H_{\text{else}}$  and therefore  $\text{rank } P' = \text{rank } \text{I} \otimes P = \text{rank } \text{I} \otimes \tilde{P}$  for small enough  $\zeta$ . Hence for all  $H_{\text{else}} \in \text{Herm}(\mathcal{H})$  we have

$$\|P'(H' + H_{\text{else}} \otimes \text{I})P' - \tilde{U}_{H_{\text{else}}}((H + H_{\text{else}}) \otimes P)\tilde{U}_{H_{\text{else}}}^\dagger\| \leq \epsilon + \zeta\|H_{\text{else}}\|. \quad (130)$$

With the above expression, we can begin steps 2 and 3. Putting  $H_{\text{else}} = 0$ , this becomes

$$\|P'H'P' - \tilde{U}(0)(H \otimes P)\tilde{U}^\dagger(0)\| \leq \epsilon, \quad (131)$$

and putting  $H_{\text{else}} = -H$  we have

$$\|P'H'P' - P'(H \otimes \text{I})P'\| \leq \epsilon + \zeta\|H\|. \quad (132)$$

Moreover, by (129) we can bound

$$\begin{aligned} \|P'(H \otimes \text{I})P' - H \otimes P\| &= \|W(\text{I} \otimes P)W^\dagger(H \otimes \text{I})W(\text{I} \otimes P)W^\dagger - H \otimes P\| \\ &= \|(W - \text{I})(\text{I} \otimes P)W^\dagger(H \otimes \text{I})W(\text{I} \otimes P)W^\dagger \\ &\quad + (\text{I} \otimes P)(W^\dagger - \text{I})(H \otimes \text{I})W(\text{I} \otimes P)W^\dagger \\ &\quad + (H \otimes P)(W - \text{I})(\text{I} \otimes P)W^\dagger \\ &\quad + (H \otimes P)(W^\dagger - \text{I})\| \\ &\leq 4\|H\| \cdot \|W - \text{I}\| \\ &\leq 8\sqrt{2}\pi K\sqrt{\zeta}\|H\|. \end{aligned} \quad (133)$$

Combining (132) and (133), we complete step 3:

$$\|P'H'P' - H \otimes P\| \leq \epsilon + (\zeta + 8\sqrt{2}\pi K\sqrt{\zeta})\|H\| := \delta. \quad (134)$$

Now we begin step 4. Let  $H^{(0)} = H \otimes P$ . Write the eigenvalues of this operator as  $\{\lambda_k\}_{k=1}^{M+1}$ , where  $\lambda_1 = 0$  and  $0 < \lambda_2 < \dots < \lambda_{M+1}$  are the  $M$  distinct eigenvalues of  $H$ . If  $H$  has any non-positive eigenvalues, then we shift both  $H$  and  $H'$  by a  $O(1)$  factor of the identity for the duration of the proof; notice that the gadget property then still holds up to a redefined  $\epsilon$  which does not affect the conclusions of this theorem. To see this, note that for  $\mu \in \mathbb{R}$ ,

$$\begin{aligned} \|P'(H' + \mu\text{I} + H_{\text{else}} \otimes \text{I})P' - \tilde{U}_{H_{\text{else}}}((H + \mu\text{I} + H_{\text{else}}) \otimes \tilde{P})\tilde{U}_{H_{\text{else}}}^\dagger\| &\leq \epsilon + \zeta\|\mu\text{I} + H_{\text{else}}\| \\ &\leq (\epsilon + \mu\zeta) + \zeta\|H_{\text{else}}\|, \end{aligned} \quad (135)$$

using (120) and absorbing  $\mu\text{I}$  into  $H_{\text{else}}$ . Hence  $H' \mapsto H' + \mu\text{I}$  and  $H \mapsto H + \mu\text{I}$  also satisfy the

assumptions of the theorem, up to replacing  $\epsilon \mapsto \epsilon + \mu\zeta$ . Moreover, if we can show that  $H' + \mu\mathbf{I}$  is an  $(\epsilon, \eta)$ -gadget for  $H + \mu\mathbf{I}$ , then applying the gadget definition shows that

$$\begin{aligned} \|P'H'P' - U(H \otimes P)U^\dagger\| &= \|P'(H' + \mu\mathbf{I})P' - U((H + \mu\mathbf{I}) \otimes P)U^\dagger\| \\ &\leq \epsilon, \end{aligned} \quad (136)$$

hence  $H'$  is an  $(\epsilon, \eta)$ -gadget for  $H$ .

We define

$$\lambda_{\text{gap}} = \min_{j \neq k} |\lambda_j - \lambda_k|, \quad (137)$$

which is  $O(1)$  since  $H$  acts on  $O(1)$  sites, and does not scale with  $\eta$  or  $\epsilon$ . Let  $\mathcal{P}_k$  be the eigenspace of  $H^{(0)}$  corresponding to  $\lambda_k$ .

We also diagonalise  $P'H'P'$  — see Fig. 10. By (131) and Weyl's inequality we can write the eigenvalues as  $\{\mu_k^{(i_k)}\}$  such that

$$|\mu_k^{(i_k)} - \lambda_k| \leq \epsilon, \quad \text{for all } i_k, \text{ and for all } k. \quad (138)$$

Let  $\mathcal{P}'_k$  be the eigenspace of  $P'H'P'$  corresponding to the eigenvalues  $\{\mu_k^{(i_k)}\}_{i_k}$ , which by (131) satisfies  $\dim \mathcal{P}'_k = \dim \mathcal{P}_k$  for  $\epsilon$  sufficiently small. Note that for  $j \neq k$  we have

$$|\mu_j^{(i_j)} - \lambda_k| \geq \lambda_{\text{gap}} - \epsilon. \quad (139)$$

We aim to construct a unitary operator which rotates all of the  $\mathcal{P}_i$  onto the  $\mathcal{P}'_i$  eigenspaces. We do this by induction, defining  $W^{(k)}$  to be a unitary operator which performs these rotations for  $i = 1, \dots, k$ . Moreover we define  $H^{(k)} = W^{(k)}H^{(0)}(W^{(k)})^\dagger$  to be the version of  $H \otimes P$  whose first  $k$  eigenspaces have been rotated in this way. We will use bounds on the direct rotation provided by Lemma 32; we will see that the direct rotations are well-defined for sufficiently small  $\zeta$  and  $\epsilon$ .

The inductive construction we use will bound the rotations by  $\|W^{(k)} - \mathbf{I}\| \leq \omega_k$ , where

$$\omega_k = \frac{\delta}{2\|H\|} \left( \left[ 1 + \frac{2\sqrt{2}\|H\|}{\lambda_{\text{gap}} - \epsilon} \right]^k - 1 \right). \quad (140)$$

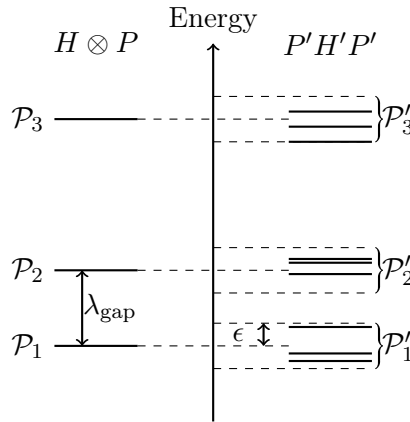


Figure 10: The degenerate eigenvalues of  $H \otimes P$  are approximated by eigenvalues of  $P'H'P'$  up to error  $\epsilon$ , corresponding to eigenspaces  $\mathcal{P}'_i$ . We construct a unitary operator which rotates the eigenspaces  $\mathcal{P}_i$  onto the  $\mathcal{P}'_i$

We now inductively define the  $H^{(k)} \in \text{Herm}(\mathcal{H} \otimes \mathcal{A})$  and  $W^{(k)} \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  as described above. For the base case, we see that clearly  $H^{(0)}$  satisfies the conditions with the trivial  $W^{(0)} = \text{I}$ .

For the inductive step, suppose we are given  $H^{(k-1)}$  and  $W^{(k-1)}$ . Notice that

$$\|P'H'P' - H^{(k-1)}\| \leq \delta + 2\omega_{k-1}\|H\| , \quad (141)$$

using (134) and the bound  $\|W^{(k-1)} - \text{I}\| \leq \omega_{k-1}$ . Hence, applying Lemma 32 to  $(P'H'P')|_{\oplus_{j \geq k} \mathcal{P}'_j}$ , and  $H^{(k-1)}|_{\oplus_{j \geq k} \mathcal{P}'_j}$  (using the fact that  $H^{(k-1)}$  is constructed to leave  $\oplus_{j \leq k-1} \mathcal{P}'_j$ , and hence also  $\oplus_{j \geq k} \mathcal{P}'_j$ , invariant), we may construct the direct rotation  $V^{(k)}$  on  $\oplus_{j \geq k} \mathcal{P}'_j$  which maps from  $W^{(k-1)}\mathcal{P}'_k(W^{(k-1)})^\dagger$  to  $\mathcal{P}'_k$ , and which satisfies

$$\begin{aligned} \|V^{(k)} - \text{I}\| &\leq \frac{\sqrt{2}}{\lambda_{\text{gap}} - \epsilon} \|(P'H'P')|_{\oplus_{j \geq k} \mathcal{P}'_j} - (H^{(k-1)})|_{\oplus_{j \geq k} \mathcal{P}'_j}\| \\ &\leq \frac{\sqrt{2}}{\lambda_{\text{gap}} - \epsilon} (\delta + 2\omega_{k-1}\|H\|) . \end{aligned} \quad (142)$$

For the above step, it is necessary to verify that the direct rotation is well-defined. If it were not, then there would be a nonzero vector  $|\psi\rangle \in W^{(k-1)}\mathcal{P}'_k(W^{(k-1)})^\dagger \cap (\mathcal{P}'_k)^\perp$ . Then, by the definition of these subspaces,

$$\langle \psi | P'H'P' | \psi \rangle \geq \lambda_{k+1} - \epsilon \geq \lambda_k + (\lambda_{\text{gap}} - \epsilon) , \quad \langle \psi | H^{(k-1)} | \psi \rangle = \lambda_k , \quad (143)$$

which would imply that

$$\|P'H'P' - H^{(k-1)}\| \geq \lambda_{\text{gap}} - \epsilon . \quad (144)$$

By (141), this is prohibited for sufficiently small  $\zeta$  and  $\epsilon$  — so we can safely assume that the direct rotation is well-defined.

Now we can let

$$W^{(k)} = (\text{I}_{\oplus_{j < k} \mathcal{P}'_j} \oplus V^{(k)})W^{(k-1)} , \quad (145)$$

and

$$H^{(k)} = W^{(k)}H^{(0)}(W^{(k)})^\dagger , \quad (146)$$

which satisfies

$$\|W^{(k)} - \text{I}\| \leq \omega_{k-1} + \frac{\sqrt{2}}{\lambda_{\text{gap}} - \epsilon} (\delta + 2\omega_{k-1}\|H\|) = \omega_k . \quad (147)$$

After  $M + 1$  inductive steps, we have constructed the operator

$$H^{(M+1)} = W^{(M+1)}(H \otimes P)(W^{(M+1)})^\dagger , \quad (148)$$

whose  $\lambda_k$ -eigenspace is  $\mathcal{P}'_k$  for all  $k$ . Hence

$$\|P'H'P' - W^{(M+1)}(H \otimes P)(W^{(M+1)})^\dagger\| \leq \epsilon . \quad (149)$$

Moreover, since  $H$  has no zero eigenvalues (since otherwise we shifted by a factor of the identity), we are guaranteed that the null space of  $H \otimes P$  is exactly that of  $(\text{I} \otimes P)$ , and  $\text{rank}(H \otimes P) = \text{rank}(\text{I} \otimes P)$ . By (129) and (131) we know that  $\text{rank}(P') = \text{rank}(\text{I} \otimes P)$  and  $\text{rank}(P'H'P') = \text{rank}(H \otimes P)$  for sufficiently small  $\zeta$  and  $\epsilon$ . So

$$\text{rank}(P') = \text{rank}(\text{I} \otimes P) = \text{rank}(H \otimes P) = \text{rank}(P'H'P') , \quad (150)$$

and the null space of  $P'H'P'$  is exactly that of  $P'$ . Hence by construction of  $W^{(M+1)}$ ,

$$P' = W^{(M+1)}(\mathbf{I} \otimes P)(W^{(M+1)})^\dagger. \quad (151)$$

We have therefore shown that  $(H', \mathcal{A})$  is a gadget as in our definition, using  $U = W^{(M+1)}$ , with accuracy  $\epsilon$  (possibly with an additional  $O(\zeta)$  if shifting of  $H'$  was necessary earlier in the proof) and  $\eta = O(\epsilon) + O(\sqrt{\zeta})$  given explicitly by

$$\eta = \omega_{M+1} = \frac{1}{2\|H\|} \left[ \left( \frac{2\sqrt{2}}{\lambda_{\text{gap}} - \epsilon} \|H\| + 1 \right)^M - 1 \right] [\epsilon + (\zeta + 8\sqrt{2}\pi K \sqrt{\zeta}) \|H\|]. \quad (152)$$

■

## D Gadget combination

In this section we prove the various gadget combination results. Below is a restatement of the general setup for the section.

### D.1 Combination of general gadgets

Here we introduce some preparatory lemmas before proving Proposition 23.

The first lemma allows us to immediately reduce gadgets to the case that the unitary  $U$  is a direct rotation, defined in Definition 30. This allows us to write  $U = e^S$  for  $S$  the generator of the direct rotation — the off-diagonal properties of  $S$  will simplify calculations considerably.

**Lemma 35.** *Suppose  $(H', \mathcal{A})$  is a  $(\eta, \epsilon)$ -gadget for  $H$ , where  $\eta < \sqrt{2}$ . Let  $\tilde{\epsilon} = \epsilon + 4\eta\|H\|$ .*

*Then  $(H', \mathcal{A})$  is also a  $(\eta, \tilde{\epsilon})$ -gadget for  $H$ , where the unitary  $U$  in the definition can be assumed to be the direct rotation  $W$ <sup>59</sup> between the subspaces defined by  $(\mathbf{I} \otimes P)$  and  $P'$ .*

*Proof of Lemma 35.* By the gadget definition we have  $U$  such that

$$P' = U(\mathbf{I} \otimes P)U^\dagger, \quad \|U - \mathbf{I}\| \leq \eta. \quad (153)$$

As shown by Davis et al.<sup>65</sup>, the direct rotation  $W$  between  $(\mathbf{I} \otimes P)$  and  $P'$  minimises  $\|W - \mathbf{I}\|$  subject to the first equality above, hence we have

$$P' = W(\mathbf{I} \otimes P)W^\dagger, \quad \|W - \mathbf{I}\| \leq \eta. \quad (154)$$

So

$$\begin{aligned} \|P'H'P' - W(H \otimes P)W^\dagger\| &\leq \epsilon + \|U(H \otimes P)U^\dagger - W(H \otimes P)W^\dagger\| \\ &\leq \epsilon + \|U(H \otimes P)U^\dagger - H \otimes P\| \\ &\quad + \|W(H \otimes P)W^\dagger - H \otimes P\| \\ &\leq \epsilon + 4\eta\|H\|. \end{aligned} \quad (155)$$

■

A gadget  $H'$  for  $H$  has  $\|P'H'P'\| \leq \|H\| + \epsilon$  by definition, however outside the span of  $P'$  there are no bounds on  $H'$ . For  $\eta$  small,  $P'$  will be close to the projector  $\mathbf{I} \otimes P$ . The following lemma provides a bound for  $H'$  when instead restricted to the span of  $\mathbf{I} \otimes P$ .

**Lemma 36.** *Suppose  $(H', \mathcal{A})$  is a  $(\eta, \epsilon)$ -gadget for  $H$ , with  $U$ ,  $P'$ , and  $P$  as in the definition, and where  $U$  is the direct rotation between  $(\mathbf{I} \otimes P)$  and  $P'$ . Assume  $\|H'\| \leq J'$ , and*

$$\|(\mathbf{I} \otimes P)H'(\mathbf{I} \otimes P^\perp)\| \leq J'_O . \quad (156)$$

Then

$$\|(\mathbf{I} \otimes P)H'(\mathbf{I} \otimes P)\| \leq \|H\| + O(\epsilon + \eta J'_O + \eta^2 J') . \quad (157)$$

*Proof of Lemma 36.* Let  $S$  be the generator of the direct rotation  $U$ , so that  $\|S\| = O(\eta)$  (see (100)) satisfies

$$P' = e^S(\mathbf{I} \otimes P)e^{-S} . \quad (158)$$

Hence

$$\|(\mathbf{I} \otimes P)H'(\mathbf{I} \otimes P)\| \leq \|(\mathbf{I} \otimes P)(e^{-S}H'e^S - H')(\mathbf{I} \otimes P)\| + \|P'H'P'\| . \quad (159)$$

By the gadget definition, we have  $\|P'H'P'\| \leq \|H\| + \epsilon$ , and using Lemma 33 we can bound the first term as

$$\|(\mathbf{I} \otimes P)(e^{-S}H'e^S - H')(\mathbf{I} \otimes P)\| \leq \|(\mathbf{I} \otimes P)[S, H'](\mathbf{I} \otimes P)\| + O(\eta^2 J') . \quad (160)$$

Furthermore, since  $S$  is off-diagonal with respect to  $(\mathbf{I} \otimes P)$  we have

$$\|(\mathbf{I} \otimes P)[S, H'](\mathbf{I} \otimes P)\| \leq O(\eta J'_O) , \quad (161)$$

completing the proof. ■

We now have the necessary tools to prove the gadget combination result Proposition 23. We are provided with the gadgets  $(H'_i, \mathcal{A}_i)$  for each of the  $H_i$  (corresponding to  $U_i, P_i, P'_i$  as in the gadget definition), which immediately suggest using  $P = \otimes_i P_i$  for the gadget  $H' = \sum_i H'_i$  for  $H = \sum_i H_i$ . It is not immediately clear what unitary  $U$ , and hence projector  $P'$ , should be used here, since the  $U_i$  do not necessarily commute so cannot be naïvely composed. The direct rotation provides a natural choice, however; writing  $U_i = e^{S_i}$  for all  $i$ , we can choose  $U = e^{\sum_i S_i}$ . The content of the proof is just a long computation to verify that this choice indeed satisfies the gadget definition.

The following proof is a generalisation of a result of Bravyi et al.<sup>31</sup>, for which which similar techniques are used.

*Proof of Proposition 23.* We begin by reducing to the case of the direct rotation. By Lemma 35, we may replace  $\epsilon$  with

$$\tilde{\epsilon} = \epsilon + 4\eta J = O(\epsilon + \eta J) , \quad (162)$$

and hence assume that each gadget  $(H'_i, \mathcal{A}_i)$  uses the direct rotation  $U_i = e^{S_i}$ . Specifically, there exists  $P_i \in \text{Proj}(\mathcal{A}_i)$  such that

$$\|(\mathbf{I} \otimes P_i)e^{-S_i}H'_ie^{S_i}(\mathbf{I} \otimes P_i) - H_i \otimes P_i\| \leq \tilde{\epsilon} , \quad (163)$$

where  $S_i$  is the generator of the direct rotation between the projectors  $\mathbf{I} \otimes P_i$  and  $P'_i := e^{S_i}(\mathbf{I} \otimes P_i)e^{-S_i}$ . We have  $\|S_i\| \leq O(\eta)$  and  $S_i$  is an anti-Hermitian operator which is block off-diagonal with respect to the projectors  $\mathbf{I} \otimes P_i$  and  $P'_i$ <sup>59</sup>.

We define the operators

$$P := \otimes_i P_i \in \text{Proj}(\otimes_i \mathcal{A}_i) , \quad S = \sum_i S_i . \quad (164)$$

We now use the triangle inequality along with Lemmas 33-34 to bound

$$\begin{aligned} & \|(\mathbf{I} \otimes P)e^{-S} H' e^S (\mathbf{I} \otimes P) - H \otimes P\| \\ & \leq \sum_i \|(\mathbf{I} \otimes P)(H'_i - [S, H'_i] + \frac{1}{2}[S, [S, H'_i]] - \frac{1}{6}[S, [S, [S, H'_i]]]) (\mathbf{I} \otimes P) \\ & \quad - H_i \otimes P\| + O(n\eta^4 J') . \end{aligned} \quad (165)$$

Now we bound the terms in the norm separately, using that  $S_j$  is block-off-diagonal with respect to  $(\mathbf{I} \otimes P_j)$ , and that  $(\mathbf{I} \otimes P_j)$  commutes with  $H'_k$  and  $S_k$  if  $j \neq k$  since  $P_j$  acts only on the ancillary  $\mathcal{A}_j$  system, whilst  $H'_k$  and  $S_k$  act on  $\mathcal{H} \otimes \mathcal{A}_k$ .

- $(\mathbf{I} \otimes P)[S, H'_i](\mathbf{I} \otimes P)$ :

Expanding  $S = \sum_j S_j$ , notice that  $(\mathbf{I} \otimes P)[S_j, H'_i](\mathbf{I} \otimes P) = 0$  whenever  $j \neq i$ , since then we can commute  $(\mathbf{I} \otimes P_j)$  past  $H'_i$ . Hence

$$(\mathbf{I} \otimes P)[S, H'_i](\mathbf{I} \otimes P) = (\mathbf{I} \otimes P)[S_i, H'_i](\mathbf{I} \otimes P) . \quad (166)$$

- $(\mathbf{I} \otimes P)[S, [S, H'_i]](\mathbf{I} \otimes P)$ :

Note that if  $j \neq k$ , then

$$(\mathbf{I} \otimes P)[S_j, [S_k, H'_i]](\mathbf{I} \otimes P) = 0 , \quad (167)$$

since at least one of them must be also different to  $i$ . Then, for instance if  $j \neq i$ , we can commute  $(\mathbf{I} \otimes P_j)$  past  $H'_i$  and  $S_k$ . Hence it remains to consider terms of the form

$$(\mathbf{I} \otimes P)[S_j, [S_j, H'_i]](\mathbf{I} \otimes P) . \quad (168)$$

In this case, if  $j \neq i$ , then we have

$$\begin{aligned} \|(\mathbf{I} \otimes P)[S_j, [S_j, H'_i]](\mathbf{I} \otimes P)\| & \leq 4\|S_j\|^2 \|(\mathbf{I} \otimes P_i)H'_i(\mathbf{I} \otimes P_i)\| \\ & \leq O(\eta^2 J + \eta^3 J'_O + \eta^4 J') , \end{aligned} \quad (169)$$

using Lemma 36, and neglecting the  $O(\eta^2 \epsilon)$  term which is dominated by the other terms in the regime of small  $\epsilon$  to which the proposition applies. Hence

$$(\mathbf{I} \otimes P)[S, [S, H'_i]](\mathbf{I} \otimes P) = (\mathbf{I} \otimes P)[S_i, [S_i, H'_i]](\mathbf{I} \otimes P) + O(\eta^2 J + \eta^3 J'_O + \eta^4 J') . \quad (170)$$

Here we have used the assumptions of bounded locality and degree in Setup 22. This ensures that only  $O(1)$  of the local terms in  $S$  appear in the commutator  $[S, H'_i]$ , and similarly for  $[S, [S, H'_i]]$ .

- $(\mathbf{I} \otimes P)[S, [S, [S, H'_i]]](\mathbf{I} \otimes P)$ :

Here we consider terms of the form  $(\mathbf{I} \otimes P)[S_j, [S_k, [S_l, H'_i]]](\mathbf{I} \otimes P)$  in various situations. Firstly, note that if none of  $j, k, l$  are equal to  $i$  then we can commute  $(\mathbf{I} \otimes P_i)$  into the commutator to obtain

$$\begin{aligned} \|(\mathbf{I} \otimes P)[S_j, [S_k, [S_l, H'_i]]](\mathbf{I} \otimes P)\| & \leq 8\|S_j\| \|S_k\| \|S_l\| \|(\mathbf{I} \otimes P_i)H'_i(\mathbf{I} \otimes P_i)\| \\ & \leq O(\eta^3 J + \eta^4 J'_O + \eta^5 J') , \end{aligned} \quad (171)$$

by Lemma 36, neglecting the  $O(\eta^3\epsilon)$  term.

If exactly two of the  $j, k, l$  are equal to  $i$  ( $k = l = i \neq j$ , say), then we can commute  $(\mathbf{I} \otimes P_j)$  past the other terms to kill the  $S_j$  term, and the expression vanishes.

If exactly one of the  $j, k, l$  is equal to  $i$ , then by commuting  $(\mathbf{I} \otimes P_i)S_i = S_i(\mathbf{I} \otimes P_i^\perp)$ , we arrive at

$$\begin{aligned} \|(\mathbf{I} \otimes P)[S_j, [S_k, [S_l, H'_i]]](\mathbf{I} \otimes P)\| &\leq O(\eta^3)\|(\mathbf{I} \otimes P_i^\perp)H'_i(\mathbf{I} \otimes P_i)\| \\ &= O(\eta^3 J'_O) . \end{aligned} \quad (172)$$

Hence

$$\begin{aligned} (\mathbf{I} \otimes P)[S, [S, [S, H'_i]]](\mathbf{I} \otimes P) &= (\mathbf{I} \otimes P)[S_i, [S_i, [S_i, H'_i]]](\mathbf{I} \otimes P) \\ &\quad + O(\eta^3 J + \eta^3 J'_O + \eta^5 J') , \end{aligned} \quad (173)$$

once again using the locality assumptions of Setup 22.

So putting the above bounds together and applying Lemma 33, we obtain

$$\begin{aligned} \|(\mathbf{I} \otimes P)e^{-S}H'e^S(\mathbf{I} \otimes P) - H \otimes P\| &\leq \sum_i \|(\mathbf{I} \otimes P)e^{-S_i}H'_ie^{S_i}(\mathbf{I} \otimes P) - H_i \otimes P\| \\ &\quad + O(n\eta^2 J + n\eta^3 J'_O + n\eta^4 J') \\ &\leq O(n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J') , \end{aligned} \quad (174)$$

where the last inequality follows because the  $H'_i$  are gadgets for the  $H_i$ .

Noting also that  $\|e^S - \mathbf{I}\| = O(n\eta)$ , this completes the proof that  $(H', \mathcal{A})$  is a  $(\eta', \epsilon')$ -gadget for  $H$  as required. ■

## D.2 Combination of low-energy gadgets

Having proved Proposition 23, we know that the  $H'$  from Proposition 24 is an  $(\eta', \epsilon')$ -gadget for  $H$ . It remains to prove that it is in fact a  $(\Delta', \eta', \epsilon')$ -gadget, which requires replacing the projector  $\tilde{P} = e^S(\mathbf{I} \otimes P)e^{-S}$  in (174) by a low-energy projector  $P_{\leq \Delta'}(H')$ . This requires the use of the following corollary to the Davis-Kahan  $\sin \theta$  theorem (Lemma 32).

**Lemma 37.** *Let  $A \in \text{Herm}(\mathcal{H})$ , and let  $P \in \text{Proj}(\mathcal{H})$  be a projector of the same rank as  $P_{\leq \Delta(A)}$ , where  $P_{\leq \Delta(A)} \in \text{Proj}(\mathcal{H})$  is the projector onto the eigenvectors of  $A$  with eigenvalues less than  $\Delta$ . Suppose that  $\|PAP\| \leq \lambda$ .*

*Then, for any  $\Delta > \lambda$ , the direct rotation  $U \in \text{U}(\mathcal{H})$  from  $P$  to  $P_{\leq \Delta(A)}$  satisfies*

$$\|U - \mathbf{I}\| \leq \frac{\sqrt{2}}{\Delta - \lambda} \|P^\perp AP\| . \quad (175)$$

*Proof of Lemma 37.* Follows from Lemma 32 using  $A \mapsto PAP$ ,  $B \mapsto A$ . ■

*Proof of Proposition 24.* By Proposition 23, we have

$$\tau := \|(I \otimes P)e^{-S}H'e^S(I \otimes P) - H \otimes P\| \leq O(n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J') , \quad (176)$$

and by condition (48) this implies that

$$\frac{\tau}{\Delta} \rightarrow 0 \quad \text{as } n \rightarrow \infty. \quad (177)$$

Letting  $\tilde{P} = e^S(I \otimes P)e^{-S}$  and applying the triangle inequality, this gives

$$\|\tilde{P}H'\tilde{P}\| \leq \|H\| + \tau . \quad (178)$$

We seek to use Lemma 37 to argue that  $\tilde{P}$  is “close to”  $P_{\leq \Delta'(H')}$ . To do this, we start by bounding  $\|\tilde{P}H'\tilde{P}^\perp\|$ . We have

$$\begin{aligned} \|\tilde{P}H'\tilde{P}^\perp\| &= \|(I \otimes P)e^{-S}H'e^S(I \otimes P^\perp)\| \\ &\leq \sum_i \|(I \otimes P)e^{-S}H'_i e^S(I \otimes P^\perp)\| , \end{aligned} \quad (179)$$

by the triangle inequality. For each  $i$ , we define  $\tilde{H}_i = e^{-S_i}H'_i e^{S_i}$ . Note that this operator is block diagonal in the basis of the projector  $(I \otimes P_i)$ , in which  $S_i$  is block off-diagonal. Here we are using the definition of the direct rotation, and the fact that  $e^{S_i}(I \otimes P_i)e^{-S_i}$  is a low-energy projector for  $H'_i$  by definition. Note also that each  $\tilde{H}_i$  acts on  $O(1)$  sites. Now we can write

$$\|\tilde{P}H'\tilde{P}^\perp\| \leq \sum_i \|(I \otimes P)e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^S(I \otimes P^\perp)\| . \quad (180)$$

From here we use Lemma 33 to expand

$$e^{S_i}\tilde{H}_i e^{-S_i} = \tilde{H}_i + [S_i, \tilde{H}_i] + R_i , \quad (181)$$

where  $R_i$  acts on  $O(1)$  sites, and by Lemma 34 we can bound  $\|R_i\| = O(J'\eta^2)$ . Hence

$$\begin{aligned} e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^S &= \tilde{H}_i + [S_i, \tilde{H}_i] + R_i \\ &\quad - [S, \tilde{H}_i + [S_i, \tilde{H}_i] + R_i] \\ &\quad + \tilde{R}_i , \end{aligned} \quad (182)$$

where the remainder  $\tilde{R}_i$  is similarly obtained by Lemmas 33-34 with  $\|\tilde{R}_i\| = O(J'\eta^2)$ . Then, using that  $S_i$ ,  $\tilde{H}_i$ , and  $R_i$  each act on  $O(1)$  sites, we can estimate

$$e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^S = \tilde{H}_i - \sum_{j \neq i} [S_j, \tilde{H}_i] + O(J'\eta^2) . \quad (183)$$



Note that only  $O(1)$  of the terms in the sum will be nonzero, and for  $j \neq i$ ,  $S_j$  will commute with  $\mathbf{I} \otimes P_i$ . Therefore

$$\begin{aligned}
\|(\mathbf{I} \otimes P)[S_j, \tilde{H}_i](\mathbf{I} \otimes P^\perp)\| &\leq 2\|S_j\| \|(\mathbf{I} \otimes P_i)\tilde{H}_i\| \\
&\leq O(\eta) \|e^{-S_i}(\mathbf{I} \otimes P_i)e^{S_i} H'_i\| \\
&= O(\eta) \|P_{\leq \Delta(H'_i)} H'_i\| \\
&\leq O(\eta)(\|H_i\| + \epsilon) = O(\eta J + \eta\epsilon) .
\end{aligned} \tag{184}$$

In the last line we have used the fact that  $H'$  defines a  $(\Delta, \eta, \epsilon)$ -gadget for  $H_i$ . Hence we can conclude that

$$(\mathbf{I} \otimes P)e^{-S} e^{S_i} \tilde{H}_i e^{-S_i} e^S (\mathbf{I} \otimes P^\perp) = O(J'\eta^2 + \eta J + \eta\epsilon) , \tag{185}$$

and so, inserting into (179), we have

$$\|\tilde{P} H' \tilde{P}^\perp\| := \omega = O(nJ'\eta^2 + n\eta J + n\eta\epsilon) . \tag{186}$$

Notice that, since we are interested in the case where  $\epsilon = o(1)$  and  $J = \theta(1)$ , the first two terms on the right-hand side will typically dominate the third one.

Now we show that the restriction of  $H'$  to the image of  $\tilde{P}^\perp$  has high-energy eigenvalues. Let  $|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle \in \mathcal{H} \otimes \mathcal{A}_i$ . We consider the expression  $(\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) H'_i (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle)$  in two cases:

- **Case 1:**  $|\psi_{\mathcal{A}_i}\rangle \in P_i \mathcal{A}_i$

Then

$$\begin{aligned}
&(\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) H'_i (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle) \\
&\geq (\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) P_{\leq \Delta(H'_i)} H'_i P_{\leq \Delta(H'_i)} (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle) \\
&\geq (\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) e^{S_i} (H_i \otimes P_i) e^{-S_i} (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle) - \epsilon \\
&\geq \langle \psi_{\mathcal{H}} | H_i | \psi_{\mathcal{H}} \rangle - (\epsilon + 2J\eta) .
\end{aligned} \tag{187}$$

- **Case 2:**  $|\psi_{\mathcal{A}_i}\rangle \in P_i^\perp \mathcal{A}_i$

Then

$$\begin{aligned}
&(\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) H'_i (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle) \\
&\geq \Delta (\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) P_{> \Delta(H'_i)} (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle) \\
&= \Delta ((\langle \psi_{\mathcal{H}} | \otimes \langle \psi_{\mathcal{A}_i} |) e^{S_i} (\mathbf{I} \otimes P_i^\perp) e^{-S_i} (|\psi_{\mathcal{H}}\rangle \otimes |\psi_{\mathcal{A}_i}\rangle)) \\
&\geq \Delta(1 - 2\eta) \\
&\geq \langle \psi_{\mathcal{H}} | H_i | \psi_{\mathcal{H}} \rangle ,
\end{aligned} \tag{188}$$

using that  $\Delta(1 - 2\eta) \geq J$  for large enough  $n$ , as  $\Delta = \Omega(nJ)$  by assumption.

Now consider any  $|\psi\rangle \in \mathcal{H} \otimes (\otimes_i \mathcal{A}_i)$  of the form

$$|\psi\rangle = e^S |\psi_{\mathcal{H}}\rangle \otimes (\otimes_i |\psi_{\mathcal{A}_i}\rangle) , \tag{189}$$

where  $P_j|\psi_{\mathcal{A}_j}\rangle = 0$  for at least one value of  $j$ . Note that such states span the image of  $\tilde{P}^\perp$ . Then

$$\begin{aligned}
\langle \psi | \tilde{P}^\perp H' \tilde{P}^\perp | \psi \rangle &= \sum_i \langle \psi | H'_i | \psi \rangle \\
&\geq \sum_{i \neq j} (\langle \psi_{\mathcal{H}} | H_i | \psi_{\mathcal{H}} \rangle - \epsilon - 2J\eta) + \Delta(1 - 2\eta) \\
&\geq \langle \psi_{\mathcal{H}} | H | \psi_{\mathcal{H}} \rangle - (N - 1)(\epsilon + 2J\eta) - J + \Delta(1 - 2\eta) \\
&\geq \Delta(1 - 2\eta) - (\|H\| + J + N(\epsilon + 2J\eta)) \\
&\geq \frac{3}{4}\Delta,
\end{aligned} \tag{190}$$

where we have used the condition (46).

Now let  $\tilde{H} := \tilde{P}H'\tilde{P} + \tilde{P}^\perp H'\tilde{P}^\perp$ , so that  $\|H' - \tilde{H}\| = O(nJ'\eta^2)$  by (186). Based on (178) and (190), we know that

$$\text{spec } \tilde{H} \subset [-(\|H\| + \tau), \|H\| + \tau] \cup [\frac{3}{4}\Delta, \infty), \tag{191}$$

corresponding to low- and high-energy projectors  $\tilde{P}$  and  $\tilde{P}^\perp$  respectively. Note that (46) in particular implies that  $\Delta \geq 4\|H\|$ , so  $\frac{1}{2}\Delta - \|H\| \geq \frac{1}{4}\Delta$ . Using this, and (177), we see

$$\frac{\frac{1}{2}\Delta - (\|H\| + \tau)}{\Delta} \geq \frac{1}{4} - \frac{\tau}{\Delta} = \Omega(1) \tag{192}$$

for large  $n$ . But by (46), (47), and (48),

$$\frac{\omega}{\Delta} = O(n\eta^2 J'/\Delta) + O(n\eta J/\Delta) + O(n\eta\epsilon/\Delta) = o(1). \tag{193}$$

for sufficiently large  $n$ , we will have  $\omega < \frac{1}{2}\Delta - (\|H\| + \tau)$ , and hence

$$\text{spec } \tilde{H} \subset (-\infty, \frac{1}{2}\Delta - \omega] \cup [\frac{3}{4}\Delta, \infty), \tag{194}$$

once again corresponding to subspaces defined by  $\tilde{P}$  and  $\tilde{P}^\perp$ . Now, by (186) and Lemma 29, we see that the full Hamiltonian  $H'$  has a ( $\leq \frac{1}{2}\Delta$ )-low energy subspace with the same dimension as the rank of  $\tilde{P}$ , and moreover

$$\text{spec } H' \subset (-\infty, \frac{1}{2}\Delta] \cup [\frac{3}{4}\Delta - \omega, \infty). \tag{195}$$

Now set  $\Delta' = \Delta/2$ . Notice that  $\|\tilde{P} - P_{\leq \Delta'}(H')\| < 1$ , since otherwise there would be a state of energy less than  $\Delta'$  in the image of  $\tilde{P}^\perp$ , which is disallowed by (191). So the direct rotation  $W \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  from  $\tilde{P}$  to  $P_{\leq \Delta'}(H')$  is well-defined, and by Lemma 37

$$\|W - \text{I}\| \leq \frac{\sqrt{2}}{\frac{1}{2}\Delta - \|H\| - \tau} \omega. \tag{196}$$

Again using that  $\Delta \geq 4\|H\|$ , so  $\frac{1}{2}\Delta - \|H\| \geq \frac{1}{4}\Delta = \Omega(\Delta)$ . By (48), this will dominate the relatively

small  $\tau$  term, so using (186) and that  $\Delta = \Omega(J')$  we have

$$\begin{aligned} \|W - \mathbf{I}\| &= O(\omega/J') = O(n\eta^2 + n\eta J/J' + n\eta\epsilon/J') \\ &= O(n\eta^2(1 + J/\eta J' + \epsilon/J'\eta)) \\ &= O\left(n\eta^2(1 + o(1/n) + o(\epsilon/nJ))\right) = O(n\eta^2), \end{aligned} \quad (197)$$

where in the last line we used (48). We can write  $W$  in terms of its anti-Hermitian and off-diagonal generator  $X$ ,  $W = e^X$ , where  $\|X\| = O(n\eta^2)$  by (100). Then

$$P_{\leq\Delta'(H')} = U(\mathbf{I} \otimes P)U^\dagger, \quad (198)$$

where  $U = e^X e^S$ . Note that

$$\begin{aligned} \|U - \mathbf{I}\| &\leq \|e^X(e^S - \mathbf{I})\| + \|e^X - \mathbf{I}\| \\ &\leq \|S\| + \|X\| \\ &= O(n\eta) + O(n\eta^2) = O(n\eta). \end{aligned} \quad (199)$$

It remains to bound find  $\epsilon'$  to achieve a bound of the form

$$\|P_{\leq\Delta'(H')}H'P_{\leq\Delta'(H')} - U(H \otimes P)U^\dagger\| \leq \epsilon'. \quad (200)$$

Using (178) and the triangle inequality we have

$$\|P_{\leq\Delta'(H')}H'P_{\leq\Delta'(H')} - U(H \otimes P)U^\dagger\| \leq \|\tilde{P}(e^{-X}H'e^X - H')\tilde{P}\| + \tau, \quad (201)$$

and the first term can be bounded using Lemma 33 (and bounding the remainder with (118)) by

$$\begin{aligned} \|\tilde{P}(e^{-X}H'e^X - H')\tilde{P}\| &\leq \|\tilde{P}[X, H']\tilde{P}\| + \frac{1}{2}\|[X, [X, H']]\| \\ &\leq 2\|X\| \cdot \|\tilde{P}H'\tilde{P}^\perp\| + 2\|X\|^2 \cdot \|H'\| \\ &\leq O(n^3\eta^4 J' + n^2\eta^3 J). \end{aligned} \quad (202)$$

In the second inequality we have used that  $X$  is off-diagonal with respect to  $\tilde{P}$ , and in the third inequality we have used (186) to bound  $\|\tilde{P}H'\tilde{P}^\perp\|$ . Hence

$$\epsilon' = O(n\epsilon + n\eta J + n\eta^3 J'_O + n^3\eta^4 J'). \quad (203)$$

■

### D.3 Ground-state estimation with bounded-strength low-energy gadgets

The following proof of Theorem 25 is a simple corollary of Proposition 23, and generalises the proof of Bravyi et al., Theorem 1<sup>31</sup>.

*Proof of Theorem 25.* For the first part of this proof, we seek to put a lower bound on the individual gadgets  $H'_i$ . We write

$$H'_i = P_{\leq\Delta(H'_i)}H'_iP_{\leq\Delta(H'_i)} + P_{>\Delta(H'_i)}H'_iP_{>\Delta(H'_i)}, \quad (204)$$

and consider the high- and low-energy parts separately.

For the low-energy part, we can apply the gadget definition to write

$$\begin{aligned} \|P_{\leq\Delta(H'_i)}H'_iP_{\leq\Delta(H'_i)} - H_i \otimes P_i\| &\leq \epsilon + \|e^{S_i}(H_i \otimes P_i)e^{-S_i} - H_i \otimes P_i\| \\ &\leq O(\epsilon + \eta J) , \end{aligned} \quad (205)$$

using Lemma 33, and hence

$$P_{\leq\Delta(H'_i)}H'_iP_{\leq\Delta(H'_i)} \geq H_i \otimes P_i + O(\epsilon + \eta J) . \quad (206)$$

For the high-energy part, we first notice that since the spectrum of  $P_{>\Delta(H'_i)}H'_iP_{>\Delta(H'_i)}$  lies in  $(\Delta, \infty)$ , where  $\Delta \geq \|H_i\| - \epsilon$  (by the assumption of the  $(\Delta, \eta, \epsilon)$ -gadget definition that  $P_{\leq\Delta(H'_i)}H'_iP_{\leq\Delta(H'_i)}$  has the same spectrum as  $H'_i$  up to error  $\epsilon$ ), and so

$$P_{>\Delta(H'_i)}H'_iP_{>\Delta(H'_i)} \geq P_{>\Delta(H'_i)}(H_i \otimes \mathbf{I})P_{>\Delta(H'_i)} + O(\epsilon) . \quad (207)$$

Furthermore, we can approximate the RHS of this expression by

$$\begin{aligned} &\|P_{>\Delta(H'_i)}(H_i \otimes \mathbf{I})P_{>\Delta(H'_i)} - H_i \otimes P_i^\perp\| \\ &= \|e^{S_i}(\mathbf{I} \otimes P_i^\perp)e^{-S_i}(H_i \otimes \mathbf{I})e^{S_i}(\mathbf{I} \otimes P_i^\perp)e^{-S_i} - H_i \otimes P_i^\perp\| \\ &\leq O(\eta J) , \end{aligned} \quad (208)$$

by applying Lemma 33. Hence

$$P_{>\Delta(H'_i)}H'_iP_{>\Delta(H'_i)} \geq H_i \otimes P_i^\perp + O(\epsilon + \eta J) . \quad (209)$$

Summing (206) and (209) for all  $i$ , we obtain

$$H' = \sum_i H'_i \geq \sum_i (H_i \otimes \mathbf{I} + O(\epsilon + \eta J)) = H \otimes \mathbf{I} + O(n\epsilon + n\eta J) , \quad (210)$$

and so the ground state energy of  $H'$  must satisfy

$$\lambda_0(H') \geq \lambda_0(H) + O(n\epsilon + n\eta J) . \quad (211)$$

Now notice that the restriction of  $H'$  to a subspace can only increase  $\lambda_0$ , so

$$\lambda_0(H') = \lambda_0(e^{-S}H'e^S) \leq \lambda_0((\mathbf{I} \otimes P)e^{-S}H'e^S(\mathbf{I} \otimes P)) . \quad (212)$$

Here, as in the work of Bravyi et al.<sup>31</sup>, we abuse notation slightly: in the expression  $\lambda_0((\mathbf{I} \otimes P)e^{-S}H'e^S(\mathbf{I} \otimes P))$ , we implicitly take the ground state of the restriction of  $e^{-S}H'e^S$  to the image of  $(\mathbf{I} \otimes P)$ . So by Proposition 23 we have

$$\lambda_0(H') \leq \lambda_0(H) + O(n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J') . \quad (213)$$

Combining (211)-(213) gives the desired result.  $\blacksquare$

## E Gadget energy scaling

In this section, we prove Theorem 26. Firstly, we introduce the notion of a  $k$ -local function, which can be thought of as a classical  $k$ -local observable on a state space  $\{0, 1\}^n$ .

**Definition 38** ( $k$ -local function). Let  $f : \{0, 1\}^n \rightarrow \mathbb{R}$  be a function. We say that  $f$  is  $k$ -local if it can be written as a sum of functions

$$f(x_1, x_2, \dots, x_n) = \sum_i f_i(x_1, x_2, \dots, x_n) , \quad (214)$$

where the  $f_i : \{0, 1\}^n \rightarrow \mathbb{R}$  each depend on at most  $k$  of their inputs.

The following simple lemmas show that there exist  $k$ -local functions which cannot be approximated well by  $k'$ -local functions for  $k' < k$ .

**Lemma 39.** Let  $f$  be a  $k$ -local function on  $n$  inputs. Then  $\mathcal{R}f : \{0, 1\}^{n-1} \rightarrow \mathbb{R}$ , defined by

$$\mathcal{R}f(x_1, \dots, x_{n-1}) = f(x_1, \dots, x_{n-1}, 0) - f(x_1, \dots, x_{n-1}, 1) \quad (215)$$

is  $(k - 1)$ -local.

*Proof of Lemma 39.* Decomposing  $f = \sum_i f_i$  as in Definition 38, note that any  $f_i$  which does not depend on  $x_n$  has  $\mathcal{R}f_i = 0$ . Moreover, any  $f_i$  which does depend on  $x_n$  depends on at most  $(k - 1)$  other inputs, hence  $\mathcal{R}f_i$  is  $(k - 1)$ -local. ■

**Lemma 40.** Let  $k > k' > 0$ . There exists a  $k$ -local function  $f : \{0, 1\}^k \rightarrow \mathbb{R}$  with  $\max_{x \in \{0, 1\}^k} |f(x)| \leq 1$  such that for any  $k'$ -local function  $g : \{0, 1\}^k \rightarrow \mathbb{R}$ ,

$$\max_{x \in \{0, 1\}^k} |f(x) - g(x)| \geq 1 . \quad (216)$$

*Proof of Lemma 40.* For any  $r \geq 1$ , we can define  $\mathcal{R}^r f : \{0, 1\}^{k-r} \rightarrow \mathbb{R}$  by

$$\mathcal{R}^r f(x_1, \dots, x_{k-r}) = \sum_{x_{k-r+1}, \dots, x_k \in \{0, 1\}} (-1)^{\sum_{j=1}^r x_{k-r+j}} f(x_1, \dots, x_k) . \quad (217)$$

Applying Lemma 39 inductively, note that  $\mathcal{R}^r f$  is  $(k - r)$ -local. In particular,  $\mathcal{R}^{k'} g$  is constant for any  $k'$ -local  $g$ .

Let  $f : \{x_1, \dots, x_k\}$  be the parity function

$$f(x_1, \dots, x_k) = (-1)^{\sum_{i=1}^k x_i} . \quad (218)$$

Then we can calculate

$$\begin{aligned} \mathcal{R}^{k'} f(x_1, \dots, x_{k-k'}) &= \sum_{x_{k-k'+1}, \dots, x_k \in \{0, 1\}} (-1)^{\sum_{j=1}^{k'} x_{k-k'+j}} \cdot (-1)^{\sum_{i=1}^k x_i} \\ &= 2^{k'} (-1)^{\sum_{j=1}^{k-k'} x_j} \end{aligned} \quad (219)$$

Hence  $\mathcal{R}^{k'} f(x_1, \dots, x_{k-k'})$  takes values  $\pm 2^{k'}$ , whereas  $\mathcal{R}^{k'} g$  is a constant function. So there must exist exist  $y \in \{0, 1\}^{k-k'}$  such that

$$|\mathcal{R}^{k'} f(y) - \mathcal{R}^{k'} g(y)| \geq 2^{k'} . \quad (220)$$

Hence, expanding  $\mathcal{R}^{k'} f(y)$  and  $\mathcal{R}^{k'} g(y)$  into  $2^{k'}$  terms using (217), we must have

$$\max_{x \in \{0,1\}^k} |f(x) - g(x)| \geq 1 . \quad (221)$$

■

Note that Lemma 40 uses the parity function (which appears in the proof of Theorem 26) for illustration, but a similar argument could apply to most  $k$ -local functions; the vector space of  $k$ -local functions has a higher dimension than that of  $k'$ -local functions.

The following proof uses the intuition from Lemma 40 to argue that the target  $k$ -local term cannot be reproduced by a  $k'$ -local Hamiltonian.

*Proof of Theorem 26.* By the gadget definition, we have  $U \in \text{U}(\mathcal{H} \otimes \mathcal{A})$  and  $P \in \text{Proj}(\mathcal{A})$  such that

$$\|U - \text{I}\| \leq \eta , \quad \|P' H' P' - U(H \otimes P)U^\dagger\| \leq \epsilon , \quad \text{where } P' = U(\text{I} \otimes P)U^\dagger . \quad (222)$$

For any given  $x \in \{0,1\}^k$ , define  $|\psi_x\rangle \in \mathcal{H}$  to be the pure state whose  $i$ th qubit is in the state  $|x_i\rangle$ . Moreover let  $|\phi\rangle \in \mathcal{A}$  be some state satisfying  $P|\phi\rangle = |\phi\rangle$ . Then define functions  $F, f : \{0,1\}^k \rightarrow \mathbb{R}$  by

$$\begin{aligned} F(x) &= \text{tr}[H|\psi_x\rangle\langle\psi_x|] , \\ f(x) &= \text{tr}[H'(|\psi_x\rangle\langle\psi_x| \otimes |\phi\rangle\langle\phi|)] . \end{aligned} \quad (223)$$

Notice that  $F$  and  $f$  are  $k$ - and  $k'$ -local respectively, and by Lemma 40 there exists some  $y \in \{0,1\}^k$  such that

$$|F(y) - f(y)| \geq J . \quad (224)$$

On the other hand, for all  $x \in \{0,1\}^k$  we have

$$\begin{aligned} |F(x) - f(x)| &= |\text{tr}[(H \otimes P)(|\psi_x\rangle\langle\psi_x| \otimes |\phi\rangle\langle\phi|)] - \text{tr}[H'(|\psi_x\rangle\langle\psi_x| \otimes |\phi\rangle\langle\phi|)]| \\ &\leq \|(\text{I} \otimes P)H'(\text{I} \otimes P) - H \otimes P\| \\ &= \|P'UH'U^\dagger P' - U(H \otimes P)U^\dagger\| \\ &\leq \epsilon + \|P'(H' - UH'U^\dagger)P'\| \\ &\leq \epsilon + 2\eta\|H'\| . \end{aligned} \quad (225)$$

Hence we must have scaling

$$\|H'\| \geq \frac{J - \epsilon}{2\eta} , \quad (226)$$

as required. ■

## F Dissipative gadgets

### F.1 Dissipative gadgets in isolation

Here we first prove Proposition 27. This follows from direct calculation, by Taylor expanding the expression  $e^{-i\delta t H'}$  and identifying the leading order terms in  $\delta t$ . This approach is complicated by the

fact that  $H'$  itself consists of terms that are  $O((\delta t)^{-1})$  and  $O((\delta t)^{-1/2})$ , but the task is simplified since we only need to calculate the time evolution of states of the form  $|\psi\rangle \otimes |0\rangle$ .

*Proof of Proposition 27.* First, notice that the requirement  $H_{|1\rangle\langle 1|}^2 = \omega^2 \mathbf{I}$  implies that

$$e^{-i\delta t H_{|1\rangle\langle 1|}} = \mathbf{I}, \quad H_{|1\rangle\langle 1|}^{-1} = \omega^{-2} H_{|1\rangle\langle 1|}. \quad (227)$$

Now we expand  $e^{-i\delta t H'}$ :

$$e^{-i\delta t H'} = \sum_{k \geq 0} \frac{(-i\delta t)^k}{k!} (H_{\mathbf{I}} \otimes \mathbf{I} + H_X \otimes X + H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1|)^k. \quad (228)$$

We can expand out this expression so that each term is a product of  $a$  factors of  $H_{\mathbf{I}} \otimes \mathbf{I}$ ,  $b$  factors of  $H_X \otimes X$ , and  $c$  factors of  $H_{|1\rangle\langle 1|}$ , for some  $a, b, c \in \mathbb{N}$ . Such a term is accompanied by  $(\delta t)^{a+b+c}$ , to give a total order of  $(\delta t)^{a+\frac{1}{2}b}$  (using that  $\|H_X\| = O((\delta t)^{-1/2})$  and  $\|H_{|1\rangle\langle 1|}\| = O((\delta t)^{-1})$ ). There are eight cases producing terms of order  $O((\delta t)^{3/2})$  and lower, which we enumerate in Fig. 11.

Case	$a$	$b$	$c$	Order
1	0	0	0	$O(1)$
2	0	0	$\mathbb{N}_+$	
3	0	1	$\mathbb{N}$	$O((\delta t)^{1/2})$
4	1	0	0	$O(\delta t)$
5	1	0	$\mathbb{N}_+$	
6	0	2	$\mathbb{N}$	
7	1	1	$\mathbb{N}$	$O(t^{3/2})$
8	0	3	$\mathbb{N}$	

Figure 11: Enumeration of possible cases for the values of  $a, b, c$  giving rise to terms of order  $O((\delta t)^{3/2})$  and lower in (228). We denote  $\mathbb{N}_+ = \{1, 2, \dots\}$  and  $\mathbb{N} = \{0, 1, 2, \dots\}$ .

In particular, we are interested in the block-elements  $(\mathbf{I} \otimes \langle 0|)e^{-i\delta t H'}(\mathbf{I} \otimes |0\rangle)$  and  $(\mathbf{I} \otimes \langle 1|)e^{-i\delta t H'}(\mathbf{I} \otimes |0\rangle)$ , since the other blocks in  $e^{-i\delta t H'}$  will annihilate states of the form  $|\psi\rangle \otimes |0\rangle$ .

- $(\mathbf{I} \otimes \langle 0|)e^{-i\delta t H'}(\mathbf{I} \otimes |0\rangle)$ :

Note that each factor of  $H_X \otimes X$  flips the ancillary qubit  $\mathcal{A}$ , whereas each factor of  $H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1|$  annihilates states with the ancillary qubit in state  $|0\rangle$ . As a result, the only contributions to  $(\mathbf{I} \otimes \langle 0|)e^{-i\delta t H'}(\mathbf{I} \otimes |0\rangle)$  from Fig. 11 are those such that  $b$  is even. Moreover,  $c$  can only be nonzero if  $b$  is at least 2 (so that the factors of  $H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1|$  can be sandwiched between two  $H_X \otimes X$  factors). This restricts us to cases 1, 4, and 6, so

$$(\mathbf{I} \otimes \langle 0|)e^{-i\delta t H'}(\mathbf{I} \otimes |0\rangle) = \mathbf{I} - i\delta t H_{\mathbf{I}} + \sum_{k \geq 2} \frac{(-i\delta t)^k}{k!} H_X H_{|1\rangle\langle 1|}^{k-2} H_X + O((\delta t)^2). \quad (229)$$

Furthermore, the sum can be simplified to

$$\begin{aligned}
\sum_{k \geq 2} \frac{(-i\delta t)^k}{k!} H_X H_{|1\rangle\langle 1|}^{k-2} H_X &= H_X H_{|1\rangle\langle 1|}^{-2} \left( \sum_{k \geq 2} \frac{(-i\delta t)^k}{k!} H_{|1\rangle\langle 1|}^k \right) H_X \\
&= H_X H_{|1\rangle\langle 1|}^{-2} \left( e^{-i\delta t H_{|1\rangle\langle 1|}} - \mathbf{I} + i\delta t H_{|1\rangle\langle 1|} \right) H_X \\
&= i\delta t \omega^{-2} H_X H_{|1\rangle\langle 1|} H_X , 
\end{aligned} \tag{230}$$

using (227). Hence we have shown that

$$(\mathbf{I} \otimes \langle 0|) e^{-i\delta t H'} (\mathbf{I} \otimes |0\rangle) = \mathbf{I} - i\delta t (H_I - \omega^{-2} H_X H_{|1\rangle\langle 1|} H_X) + O((\delta t)^2) . \tag{231}$$

- $(\mathbf{I} \otimes \langle 1|) e^{-i\delta t H'} (\mathbf{I} \otimes |0\rangle)$ :

By a similar argument to above, the only contributing terms from Fig. 11 are those such that  $b$  is odd, so we can reduce to the cases 3, 7, and 8. In case 3, also notice that the  $H_X \otimes X$  term must appear on the right of all the  $H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1|$  terms. Hence

$$(\mathbf{I} \otimes \langle 1|) e^{-i\delta t H'} (\mathbf{I} \otimes |0\rangle) = \sum_{k \geq 1} \frac{(-i\delta t)^k}{k!} H_{|1\rangle\langle 1|}^{k-1} H_X + O((\delta t)^{3/2}) . \tag{232}$$

The sum here can be similarly simplified by (227):

$$\begin{aligned}
\sum_{k \geq 1} \frac{(-i\delta t)^k}{k!} H_{|1\rangle\langle 1|}^{k-1} H_X &= H_{|1\rangle\langle 1|}^{-1} \left( \sum_{k \geq 0} \frac{(-i\delta t)^k}{k!} H_{|1\rangle\langle 1|}^k - \mathbf{I} \right) H_X \\
&= H_{|1\rangle\langle 1|}^{-1} (e^{-i\delta t H_{|1\rangle\langle 1|}} - \mathbf{I}) H_X \\
&= 0 , 
\end{aligned} \tag{233}$$

so

$$(\mathbf{I} \otimes \langle 1|) e^{-i\delta t H'} (\mathbf{I} \otimes |0\rangle) = O((\delta t)^{3/2}) , \tag{234}$$

which, along with (231), completes the proof. ■

## F.2 Trotter errors

The idea for the proof of Proposition 28 is to factorise the overall evolution operator

$$e^{-i\delta t (H' + H_{\text{else}} \otimes \mathbf{I})} \approx e^{-i\delta t H_{\text{else}} \otimes \mathbf{I}} e^{-i\delta t H'} . \tag{235}$$

From here, we simply apply Proposition 27 to the initial state  $|\psi\rangle \otimes |0\rangle$ , and then evolve the  $\mathcal{H}$  system of the resultant state under  $H_{\text{else}}$ . The technical difficulty is in bounding the errors of this Trotter expansion in a way which does not depend on the size of the system. Qualitatively, one might expect this behaviour due to the bounded spread of correlations in the system over a short time  $\delta t$ , under which only a limited set of interactions in  $H_{\text{else}}$  can “interfere” with the evolution under  $H'$ . The difficulty of obtaining such bounds is compounded by the presence of terms in  $H'$  which scale as  $O((\delta t)^{-1})$  and  $O((\delta t)^{-1/2})$ . Our approach uses an explicit form of the Trotter error given by Childs et al. <sup>66</sup>. We briefly outline this process here.



Let  $A, B \in \text{Lin}(\mathcal{H})$ . We aim to find an expression for the Trotter error incurred by the expansion  $e^{t(A+B)} \approx e^{tA}e^{tB}$ .

Observe that the function  $f(t) = e^{tA}e^{tB}$  satisfies the differential equation

$$\begin{aligned} f'(t) &= Ae^{tA}e^{tB} + e^{tA}Be^{tB} \\ &= (A + B)f(t) + e^{tA}(B - e^{-tA}Be^{tA})e^{tB} . \end{aligned} \quad (236)$$

This differential equation, with initial condition  $f(0) = \text{I}$ , can be solved using following lemma.

**Lemma 41** (Variation of parameters formula<sup>66</sup>). *Let  $K \in \text{Lin}(\mathcal{H})$ , and let  $L(t) \in \text{Lin}(\mathcal{H})$  be a continuous operator-valued function of  $t$ . Suppose that  $f(t)$  satisfies the differential equation*

$$f'(t) = Kf(t) + L(t) , \quad f(0) = \text{I} . \quad (237)$$

Then there is a unique solution for  $f$  which is given by

$$f(t) = e^{tK} + \int_0^t d\tau e^{(t-\tau)K}L(\tau) . \quad (238)$$

Hence, using Lemma 41 with  $K = A + B$  and  $L(t) = e^{tA}(B - e^{-tA}Be^{tA})e^{tB}$ , we find that the Trotter error is given by

$$e^{tA}e^{tB} - e^{t(A+B)} = \int_0^t d\tau e^{(t-\tau)(A+B)}e^{\tau A}(B - e^{-\tau A}Be^{\tau A})e^{\tau B} \quad (239)$$

The expression (239) is particularly convenient because, when  $A$  is a local Hamiltonian and  $B$  acts only on  $O(1)$  sites, the bracketed term  $(B - e^{-\tau A}Be^{\tau A})$  can be bounded independently of  $n$ .

**Lemma 42.** *Let  $H = \sum_i h_i$  be a  $k$ -local Hamiltonian on a system  $\mathcal{H} = \otimes_{i=1}^n \mathcal{H}_i$ , with the degree of the interaction hypergraph bounded by an  $O(1)$  constant and  $\|h_i\| = O(1)$ . Let  $A$  be an observable supported on a set of  $O(1)$  sites. Then*

$$\|e^{itH}Ae^{-itH} - A\| \leq O(\|A\|t) . \quad (240)$$

*Proof of Lemma 42.* For  $X \in \text{Lin}(\mathcal{H})$ , define  $f_X(t) = \text{tr}[Xe^{itH}Ae^{-itH} - A]$ , so that

$$\|e^{itH}Ae^{-itH} - A\| = \max_{X \in D(\mathcal{H})} |f_X(t)| . \quad (241)$$

We can see that  $f_X(0) = 0$  and  $f'_X(t) = i \text{tr}[Xe^{itH}[H, A]e^{-itH}]$ . Moreover, since  $H$  is local on an  $O(1)$ -degree hypergraph, and  $A$  is supported on an  $O(1)$  set, only  $O(1)$  terms in  $H$  contribute to the commutator and hence  $|f'_X(t)| \leq \|X\|_1\|[H, A]\| = O(\|X\|_1\|A\|)$ . By the mean value theorem, we therefore deduce that

$$f_X(t) = O(\|X\|_1\|A\|t) , \quad (242)$$

so by (241) we are done. ■

### F.3 Dissipative gadgets with other terms

*Proof of Proposition 28.* Using (239) with  $t = \delta t$ ,  $A = -iH_{\text{else}} \otimes \text{I}$ , and  $B = -iH'$ , we obtain a Trotter

error given by

$$\begin{aligned} e^{-i\delta t H_{\text{else}} \otimes \mathbb{I}} e^{-i\delta t H'} - e^{-i\delta t (H_{\text{else}} \otimes \mathbb{I} + H')} &:= E \\ &= -i \int_0^{\delta t} d\tau e^{-i(\delta t - \tau)(H_{\text{else}} \otimes \mathbb{I} + H')} e^{-i\tau H_{\text{else}} \otimes \mathbb{I}} (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) e^{-i\tau H'} . \end{aligned} \quad (243)$$

We can write  $E$  in block form in the basis of the ancillary space,

$$E = \begin{pmatrix} (\mathbb{I} \otimes \langle 0|) E (\mathbb{I} \otimes |0\rangle) & (\mathbb{I} \otimes \langle 0|) E (\mathbb{I} \otimes |1\rangle) \\ (\mathbb{I} \otimes \langle 1|) E (\mathbb{I} \otimes |0\rangle) & (\mathbb{I} \otimes \langle 1|) E (\mathbb{I} \otimes |1\rangle) \end{pmatrix} , \quad (244)$$

and we will focus on individually bounding these blocks. Notice that, commuting projectors on the ancillary space past  $H_{\text{else}} \otimes \mathbb{I}$  and applying Lemma 42, we have

$$\begin{aligned} &\|(\mathbb{I} \otimes \langle 0|)(H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}})(\mathbb{I} \otimes |0\rangle)\| \\ &= \|(\mathbb{I} \otimes \langle 0|)H'(\mathbb{I} \otimes |0\rangle) - e^{i\tau H_{\text{else}}}(\mathbb{I} \otimes \langle 0|)H'(\mathbb{I} \otimes |0\rangle)e^{-i\tau H_{\text{else}}}\| \\ &= \|H_{\mathbb{I}} - e^{i\tau H_{\text{else}}} H_{\mathbb{I}} e^{-i\tau H_{\text{else}}}\| = O(\delta t) . \end{aligned} \quad (245)$$

With a similar process for the other blocks, we obtain the following bounds:

$$\|(\mathbb{I} \otimes \langle 0|)(H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}})(\mathbb{I} \otimes |0\rangle)\| = O(\delta t) , \quad (246a)$$

$$\|(\mathbb{I} \otimes \langle 1|)(H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}})(\mathbb{I} \otimes |0\rangle)\| = O(\delta t^{1/2}) , \quad (246b)$$

$$\|(\mathbb{I} \otimes \langle 0|)(H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}})(\mathbb{I} \otimes |1\rangle)\| = O(\delta t^{1/2}) , \quad (246c)$$

$$\|(\mathbb{I} \otimes \langle 1|)(H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}})(\mathbb{I} \otimes |1\rangle)\| = O(1) . \quad (246d)$$

Since our initial state is of the form  $|\psi\rangle \otimes |0\rangle$ , we need only bound the magnitudes of the blocks  $(\mathbb{I} \otimes \langle 0|)E(\mathbb{I} \otimes |0\rangle)$  and  $(\mathbb{I} \otimes \langle 1|)E(\mathbb{I} \otimes |0\rangle)$ . To this end, we need to describe the action of the operator  $e^{-i\tau H'}$  on the operators  $(\mathbb{I} \otimes |0\rangle)$  and  $(\mathbb{I} \otimes |1\rangle)$ , for  $0 \leq \tau \leq \delta t$ . By considering the series expansion as in (228), and noting that the ancillary qubit can only be flipped by a  $H_X \otimes X$  term of order  $O((\delta t)^{1/2})$ , we see that

$$e^{-i\tau H'}(\mathbb{I} \otimes |0\rangle) = O(1) \otimes |0\rangle + O((\delta t)^{1/2}) \otimes |1\rangle , \quad (247a)$$

$$e^{-i\tau H'}(\mathbb{I} \otimes |1\rangle) = O((\delta t)^{1/2}) \otimes |0\rangle + O(1) \otimes |1\rangle . \quad (247b)$$

Notice that here we abuse big-O notation for matrices; for example, in the above expression  $O((\delta t)^{1/2})$  should be interpreted as a matrix with operator norm bounded by  $O(\delta t^{1/2})$ . We can also crudely upper bound  $E$  as follows:

$$\begin{aligned} \|E\| &\leq \int_0^{\delta t} d\tau \|H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}\| \\ &\leq \int_0^{\delta t} d\tau O(1) = O(\delta t) , \end{aligned} \quad (248)$$

using (246d) to give the most pessimistic bound. Therefore in particular

$$e^{i\tau H_{\text{else}} \otimes \mathbb{I}} e^{i(\delta t - \tau)(H_{\text{else}} \otimes \mathbb{I} + H')} = e^{i\delta t H_{\text{else}} \otimes \mathbb{I}} e^{i(\delta t - \tau)H'} + O(\delta t) , \quad (249)$$

so, using (247a),

$$\begin{aligned} e^{i\tau H_{\text{else}} \otimes \mathbb{I}} e^{i(\delta t - \tau)(H_{\text{else}} \otimes \mathbb{I} + H')} (\mathbb{I} \otimes |0\rangle) &= e^{-i\delta t H_{\text{else}} \otimes \mathbb{I}} e^{i(\delta t - \tau)H'} (\mathbb{I} \otimes |0\rangle) + O(\delta t) \\ &= O(1) \otimes |0\rangle + O((\delta t)^{1/2}) \otimes |1\rangle, \end{aligned} \quad (250)$$

and similarly

$$e^{i\tau H_{\text{else}} \otimes \mathbb{I}} e^{i(\delta t - \tau)(H_{\text{else}} \otimes \mathbb{I} + H')} (\mathbb{I} \otimes |1\rangle) = O((\delta t)^{1/2}) \otimes |0\rangle + O(1) \otimes |1\rangle. \quad (251)$$

We can now obtain the necessary bounds on the blocks of  $E$ . Firstly, we have

$$\begin{aligned} &\|(\mathbb{I} \otimes \langle 0|) E (\mathbb{I} \otimes |0\rangle)\| \\ &\leq \int_0^{\delta t} d\tau \|(\mathbb{I} \otimes \langle 0|) e^{-i(\delta t - \tau)(H_{\text{else}} \otimes \mathbb{I} + H')} e^{-i\tau H_{\text{else}} \otimes \mathbb{I}} (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) e^{-i\tau H'} (\mathbb{I} \otimes |0\rangle)\| \\ &\leq \int_0^{\delta t} d\tau \left[ O(1) \|(\mathbb{I} \otimes \langle 0|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |0\rangle)\| \right. \\ &\quad + O((\delta t)^{1/2}) \|(\mathbb{I} \otimes \langle 1|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |0\rangle)\| \\ &\quad + O((\delta t)^{1/2}) \|(\mathbb{I} \otimes \langle 0|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |1\rangle)\| \\ &\quad \left. + O(\delta t) \|(\mathbb{I} \otimes \langle 1|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |1\rangle)\| \right] \\ &= O((\delta t)^2), \end{aligned} \quad (252)$$

using (246a-246d). Similarly, we can bound

$$\begin{aligned} &\|(\mathbb{I} \otimes \langle 1|) E (\mathbb{I} \otimes |0\rangle)\| \\ &\leq \int_0^{\delta t} d\tau \|(\mathbb{I} \otimes \langle 1|) e^{-i(\delta t - \tau)(H_{\text{else}} \otimes \mathbb{I} + H')} e^{-i\tau H_{\text{else}} \otimes \mathbb{I}} (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) e^{-i\tau H'} (\mathbb{I} \otimes |0\rangle)\| \\ &\leq \int_0^{\delta t} d\tau \left[ O((\delta t)^{1/2}) \|(\mathbb{I} \otimes \langle 0|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |0\rangle)\| \right. \\ &\quad + O(1) \|(\mathbb{I} \otimes \langle 1|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |0\rangle)\| \\ &\quad + O(\delta t) \|(\mathbb{I} \otimes \langle 0|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |1\rangle)\| \\ &\quad \left. + O((\delta t)^{1/2}) \|(\mathbb{I} \otimes \langle 1|) (H' - e^{i\tau H_{\text{else}} \otimes \mathbb{I}} H' e^{-i\tau H_{\text{else}} \otimes \mathbb{I}}) (\mathbb{I} \otimes |1\rangle)\| \right] \\ &= O((\delta t)^{3/2}). \end{aligned} \quad (253)$$

With the bounds (252) and (253) on the blocks of the Trotter error we can now conclude that

$$\begin{aligned} e^{-i\delta t(H' + H_{\text{else}} \otimes \mathbb{I})} (|\psi\rangle \otimes |0\rangle) &= e^{-i\delta t H_{\text{else}} \otimes \mathbb{I}} e^{-i\delta t H'} (|\psi\rangle \otimes |0\rangle) + E(|\psi\rangle \otimes |0\rangle) \\ &= e^{-i\delta t H_{\text{else}} \otimes \mathbb{I}} (e^{-i\delta t H} |\psi\rangle + O((\delta t)^2) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle) \\ &\quad + O((\delta t)^2) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle \\ &= (e^{-i\delta t H_{\text{else}}} e^{-i\delta t H} |\psi\rangle + O((\delta t)^2) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle), \end{aligned} \quad (254)$$

where in the second inequality we invoke Proposition 27. It remains only to bound the Trotter error in the product  $e^{-i\delta t H_{\text{else}}} e^{-i\delta t H}$ , which we can accomplish similarly. Using (239) with  $t = \delta t$ ,  $A = -iH_{\text{else}}$ ,

$B = -iH$ , we obtain

$$\begin{aligned}
& e^{-i\delta t H_{\text{else}}} e^{-i\delta t H} - e^{-i\delta t (H + H_{\text{else}})} \\
&= -i \int_0^{\delta t} d\tau e^{-i(\delta t - \tau)(H + H_{\text{else}})} e^{-i\tau H_{\text{else}}} (H - e^{i\tau H_{\text{else}}} H e^{-i\tau H_{\text{else}}}) e^{-i\tau H} .
\end{aligned} \tag{255}$$

So by Lemma 42 we have

$$e^{-i\delta t H_{\text{else}}} e^{-i\delta t H} - e^{-i\delta t (H + H_{\text{else}})} \leq O((\delta t)^2) \tag{256}$$

Combining (256) with (254) completes the proof. ■