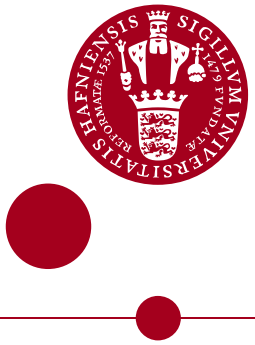
UNIVERSITY OF COPENHAGEN
FACULTY OF SCIENCE



PhD thesis

Analogue quantum simulation and perturbation theory

Dylan Harley

Advisor: Matthias Christandl, University of Copenhagen

Submitted: June 12, 2026

This thesis has been submitted to the PhD School of The Faculty of Science, University of Copenhagen

Abstract

Analogue quantum simulation is the general technique of encoding the behaviour of one physical system of interest into a different simulator system which is experimentally accessible. In this way, the system of interest can be indirectly studied by observing the natural evolution of a artificial simulator system. This leads to a natural yet subtle question: when can one family of Hamiltonians simulate another? This question has been studied from two directions: the experimental perspective, in order to characterise the expressive power of physically realisable devices; and the complexity-theoretic perspective, where the hardness of problems in physics and chemistry can be related through the simulation of hard instances. In the study of the latter perspective, a wide toolbox of techniques involving perturbation theory have been developed and applied to show that very simple families of Hamiltonian (restricted in both geometry and the types of interactions they contain) suffice to simulate all others. These conclusions are not easy to connect to the former perspective, however: perturbative simulations generically require enormous interaction strengths which are experimentally unrealistic.

In this thesis, we examine this tension and explore the conditions under which these unrealistic features are necessary for analogue simulations, and when they can be avoided by working outside of the usual perturbation theory framework. We first do this by generalising previous notions of local Hamiltonian simulation and establishing a no-go result which rules out locality reduction without strong interactions. We also describe a way to circumvent this theorem using dissipative dynamics. In the final part of this thesis, we show how classical postprocessing techniques can be used to reduce the overhead of analogue simulators by extrapolating the properties of inaccurate simulations with weaker interactions to the perturbative limit.

Resumé

Analog kvantesimulering er den generelle teknik til at indkode adfærden af et fysisk system af interesse i et andet simulatorsystem, som er eksperimentelt tilgængeligt. På denne måde kan systemet af interesse studeres indirekte ved at observere den naturlige udvikling af et kunstigt simulatorsystem. Dette leder til et naturligt, men subtilt spørgsmål: hvornår kan en familie af Hamiltonoperatorer simulere en anden? Dette spørgsmål er blevet undersøgt fra to perspektiver: det eksperimentelle perspektiv, med henblik på at karakterisere den ekspressive styrke af fysisk realiserbare apparater; og det kompleksitetsteoretiske perspektiv, hvor sværhedsgraden af problemer inden for fysik og kemi kan relateres gennem simulering af svære instanser. I studiet af det sidstnævnte perspektiv er der udviklet og anvendt en bred værktøjskasse af teknikker baseret på perturbationsteori til at vise, at meget simple familier af Hamiltonoperatorer (begrænset i både geometri og de typer interaktioner, de indeholder) er tilstrækkelige til at simulere alle andre. Disse konklusioner er dog ikke nemme at forbinde til det førstnævnte perspektiv: perturbative simuleringer kræver generelt enorme interaktionsstyrker, som er eksperimentelt urealistiske.

I denne afhandling undersøger vi denne spænding og udforsker de betingelser, hvorunder disse urealistiske egenskaber er nødvendige for analoge simuleringer, og hvornår de kan undgås ved at arbejde uden for den sædvanlige perturbationsteoretiske ramme. Vi gør dette først ved at generalisere tidligere begreber om lokal Hamiltonsimulering og etablere et no-go-resultat, som udelukker lokalitetsreduktion uden stærke interaktioner. Vi beskriver også en måde at omgå denne sætning ved hjælp af dissipativ dynamik. I den sidste del af denne afhandling viser vi, hvordan klassiske efterbehandlingsteknikker kan bruges til at reducere overheadet for analoge simulatorer ved at ekstrapolere egenskaberne af upræcise simuleringer med svagere interaktioner til den perturbative grænse.

Contributions and structure

In Section 1, we present a general introduction to the theory of quantum simulation, and motivate the study of analogue quantum simulation as an approach for near-term experiments. In Section 2, we give a complexity-theoretic motivation for the study of analogue quantum simulation and review some simulation tasks, known hardness results, and open questions. Section 3 is a lightly edited version of Ref. [1] (which also appeared in the Author's MSc thesis as part of an integrated programme), which was joint work with Ishaun Datta, Frederik Ravn Klausen, Andreas Bluhm, Daniel Stilck França, Albert Werner, and Matthias Christandl:

> [1]: Dylan Harley, Ishaun Datta, Frederik Ravn Klausen, Andreas Bluhm, Daniel Stilck França, Albert H. Werner, and Matthias Christandl. Going beyond gadgets: The importance of scalability for analogue quantum simulators. *Nature Communications*, 15(1):6527, 2024.

Section 4 is a lightly edited version of Ref. [2], which was joint work with Matthias Christandl:

> [2]: Dylan Harley and Matthias Christandl. Analogue quantum simulation with polylogarithmic interaction strengths by extrapolating within phases of matter. *arXiv preprint arXiv:2605.11285*, 2026.

The following papers were also co-authored over the course of the PhD, but do not form part of this thesis:

> [3]: Mihael Erakovic, Freek Witteveen, Dylan Harley, Jakob Günther, Moritz Bensberg, Oinam Romesh Meitei, Minsik Cho, Troy Van Voorhis, Markus Reiher, and Matthias Christandl. High ground state overlap via quantum embedding methods. *PRX Life*, 3(1):013003, 2025.

> [4]: Dylan Harley, Freek Witteveen, and Daniel Malz. Computational complexity of injective projective entangled pair states. *arXiv preprint arXiv:2509.19963*, 2025.

> [5]: Dylan Harley and Robert Koenig. Fault-tolerant one-shot entanglement generation with constant-sized quantum devices in the plane. *arXiv preprint arXiv:2604.05870*, 2026.

Acknowledgements

First of all I would like to thank my supervisor, Matthias Christandl, for his guidance and support throughout this PhD programme. It has been a fun and scientifically stimulating few years, and I am very grateful for the excellent opportunities I have had to develop and gain confidence as a researcher.

I thank Robert König for hosting me on my change of research environment, and for taking the time for many interesting discussions and the enjoyable collaboration that resulted. I am also grateful to the rest of the group at TUM for being so consistently friendly and welcoming throughout my visits.

I also wish to express gratitude to all of the other wonderful collaborators whom I have had the pleasure to work with and learn from over the past four years, with particular thanks to Freek Witteveen and Daniel Malz.

I thank all of my colleagues at QMATH, past and present, for the wonderful social and scientific environment of which I have been so privileged to be a part. The attendance rate at after-work runs was often satisfactory, for which I thank and commend all participants. Thank you to Jakob Günther and Tore Friis for proof-reading parts of this thesis and thus implicitly taking on full responsibility for all remaning typos.

Contents

1 Introduction

1.1 The computational cost of living in a quantum universe

Among the many clichés of popular physics literature are several variations on the following quote of Richard Feynman [6]:

> *"I think I can safely say that nobody understands quantum mechanics."*

It is convenient to begin a thesis in this way because it provides the Author with licence to subsequently write something completely incomprehensible, but it is also true in at least two concrete respects:

1. Quantum mechanics is difficult to understand for humans, who typically exist on length scales from $\sim 10^{-1}$ m upon initialisation up to $\sim 10^{0}$ m in their steady state. Phenomena at such scales are very well-described by Newtonian physics, whereas quantum effects such as wave-particle duality, superposition, and destructive measurement only become noticeable at the scale of atoms and molecules. This is well beyond the regime of ordinary human experience, and information about these phenomena can only be obtained from implausible-looking equations and the baffling claims of experimental physicists.

2. In fact, the mathematical structure of quantum mechanics makes it difficult to understand even for computers, compared to classical physics. The most immediate way to see this is by attempting to write down the state of a system of N particles. Classically, this is easy: we can simply write the states of all N particles one by one in a list — this takes up an amount of space scaling linearly with N. In contrast, for an analogous quantum system one must account for all possible configurations existing in superposition; generally the number of parameters required to describe such a mixture grows exponentially in N, and therefore the task quickly becomes hopeless as N increases. All of these parameters must be taken into account whilst computing how such a system evolves in time, or solving typical problems such as determining its minimum-energy configuration.

Point 1 can be marketed as an exciting feature of our universe, resolving many scientific mysteries of the early twentieth century (Why don't electrons fall into nuclei? Why does the spectrum of black body radiation contradict classical predictions? Why do small particles sometimes behave like waves?), and giving rise to many surprising and theoretically interesting phenomena beyond classical physics from the EPR paradox [7, 8], to superconductivity [9], and the vast toolbox of quantum information theory [10]. Meanwhile, Point 2 appears to pose a fundamental obstruction to solving any physical problems in practice: the behaviour of modestly sized quantum systems is out of reach for classical computation, even at the scale of supercomputers. This is a major challenge to overcome, and offers a fairly convincing discouragement from bothering to research computational many-body quantum mechanics at all. However, for better or worse, since its introduction in the early twentieth century quantum mechanics has become the

basis for much of modern science, and an understanding of its effects at the microscopic scale is essential across a range of domains. For chemistry, quantum mechanics governs the behaviour of electrons in molecules and determines reaction pathways [11]; on larger scales, this provides information useful for molecular biology [12]; in theoretical physics, quantum mechanics is necessary for the study of materials science [13] and high-energy physics [14].

The key idea to approach these challenges, first formulated almost fifty years ago independently by Manin and Feynman [15, 16], is to delegate the task to experimentalists and instead work in a computational model in which the hardware itself is fundamentally quantum. In this way, the problem of classically representing quantum superpositions can be circumvented by instead preparing and manipulating quantum states directly on a programmable quantum device. The concept of computing with quantum devices has sparked decades of subsequent theoretical research, particularly accelerating after several works which established that (a) quantum computation might even be useful for problems which are not obviously quantum [17, 18, 19], and (b) long quantum computations can be protected from noise by the use of quantum error correction [20, 21].

The engineering task implicit in the study of quantum computing is rather challenging, and no quantum computer capable of solving useful problems beyond the reach of classical computers has yet been built by man or beast (less practical computations notwithstanding [22, 23, 24, 25]). Progress has accelerated over recent years however, with numerous plucky experimental groups making remarkable progress towards the theoretical vision of a fault-tolerant programmable quantum device [26, 27, 24, 28, 29] — nowadays, an experimental demonstration of useful quantum computation appears inevitable.

This thesis will focus on the original motivating problem for quantum technologies: the simulation of quantum systems. Decades on from the initial visions of Manin and Feynman, simulation of physics and chemistry remains a leading candidate for useful quantum advantage in both the near and far term [30, 31, 32].

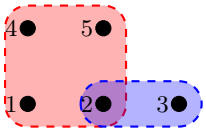
1.2 Digital and analogue quantum simulation

In this section, we will introduce some more technical terminology, and describe the basics of quantum simulation with quantum computers. The subsequent parts of this thesis will mostly focus on analogue quantum simulation.

Terminology and notation

We only consider finite-dimensional systems in this thesis, typically systems of $n \in \mathbb{N}$ qubits labelled by an index $i \in [n] := \{1, \ldots, n\}$. The state space of a single qubit is described by the Hilbert space $\mathcal{H}_i \cong \mathbb{C}^2$ (most of the results in this thesis are also valid for qudits $\mathbb{C}^d$ for $d > 2$), and the composite system is given by the tensor product

$$\mathcal{H} := \bigotimes_{i=1}^{n} \mathcal{H}_i \ . \tag{1.1}$$



$$H = H_{1245} + H_{23}$$

Figure 1.1: The interaction hypergraph of a 4-local Hamiltonian H on five sites, consisting of a 4-local term H_{1245} and a 2-local term H_{23}.

We use bra-ket notation to denote pure states (elements $|\psi\rangle \in \mathcal{H}$) and their conjugates (elements $\langle\psi|$ of the dual space $\mathcal{H}^*$). General mixed states are positive semidefinite matrices $\rho \in \mathrm{PSD}(\mathcal{H})$ such that $\mathrm{tr}[\rho] = 1$, which can be equivalently interpreted as probabilistic mixtures $\rho = \sum_x p_x |\psi_x\rangle\langle\psi_x|$ where $\{p_x\}_x$ is a probability distribution and $\{|\psi_x\rangle\}_x$ are pure states.

Physically observable quantities are represented by Hermitian matrices $O \in \mathrm{Herm}(\mathcal{H})$. In general, the measurement of such an observable on a state ρ is non-deterministic, but its expectation value is given by $\mathrm{tr}[O\rho]$. If O has the form $O = O_A \otimes \mathbb{1}_{[n]\setminus A}$ for $A \subseteq [n]$, where O_A is an observable on $\mathcal{H}_A := \bigotimes_{i\in A} \mathcal{H}_i$ and $\mathbb{1}_{[n]\setminus A}$ is the identity on all sites $[n] \setminus A$, then we say that O is supported on A. Physically, this means that a measurement of O only depends on the marginal state of ρ on the sites A, $\rho_A := \mathrm{tr}_{[n]\setminus A}[\rho]$. Often we suppress factors of the identity and just write O_A in place of O.

One observable of particular interest is the Hamiltonian $H \in \mathrm{Herm}(\mathcal{H})$, which defines the energy of the system and generates its time evolution. Typically, H is constructed as a sum of many locally supported terms $H = \sum_{A\subseteq[n]} H_A$, where the individual H_A are referred to as interactions. If $H_A = 0$ for $|A| > k$ (so that H only consists of interactions between k or fewer sites), we say that H is k-local. It is useful to represent H like this for several reasons: firstly, it explicitly shows which sites directly interact. This can be visualised via the interaction (hyper)graph of the Hamiltonian, with vertex set $[n]$ and a (hyper)edge $A \subseteq [n]$ at any subset A where $H_A \neq 0$. In physical systems, typically interaction only occurs between spatially nearby sites such as nearest neighbours on a D-dimensional lattice $\Gamma \subseteq \mathbb{Z}^D$. We will use the phrase "geometrically local" to describe this situation. Secondly, decomposition into k-local terms allows an efficient (polynomial-length) representation, since there are at most $\binom{n}{k} \sim n^k$ terms H_A which are each constant-sized $(2^k \times 2^k)$ matrices. As a result, a k-local Hamiltonian is parametrised by $\sim \mathrm{poly}(n)$ complex numbers, as opposed to $\sim \exp(\Omega(n))$ for a general element of $\mathrm{Herm}(\mathcal{H})$. Efficient descriptions are necessary for complexity theory and for classical and quantum simulation algorithms in practice. An example 4-local Hamiltonian, with associated interaction hypergraph, is depicted in Fig. 1.1.

The Hamiltonian governs the time evolution of a quantum system via the Schrödinger equation: pure states $|\psi(t)\rangle$ evolve in time as

$$\frac{\mathrm{d}}{\mathrm{d}t}|\psi(t)\rangle = -iH|\psi(t)\rangle \ , \tag{1.2}$$

and correspondingly mixed states $\rho(t)$ evolve as

$$\frac{\mathrm{d}}{\mathrm{d}t}\rho(t) = i[H, \rho(t)] \ . \tag{1.3}$$

If H is time-independent then these equations can be solved in terms of matrix exponentials:

$$|\psi(t)\rangle = e^{-itH}|\psi(0)\rangle \ , \quad \rho(t) = e^{-itH}\rho(0)e^{itH} \ . \tag{1.4}$$

In principle this gives a complete description of the behaviour of the system at arbitrary times, but the compact notation hides an impossibly complex computational task: H is a $2^n \times 2^n$ matrix which must be exponentiated and applied to the (similarly exponential-sized) state vector or density matrix.

Broadly speaking, there are two natural classes of computational problems which we consider.

- **Dynamical properties.** Combining the above notions, we may ask the general question: given a Hamiltonian H, an initial state $\rho(0)$, an observable O, and a time $t \geq 0$, what is the expected value of the observable O after time t? Equivalently, this question boils down to computing values of the function $f(t)$ defined by

  $$f(t) := \mathrm{tr}\left[O\rho(t)\right] = \mathrm{tr}\left[Oe^{-itH}\rho(0)e^{-itH}\right] \ . \tag{1.5}$$

  Such questions have obvious physical motivation: this is exactly the calculation that Nature must perform every time an experimentalist takes a measurement. We can therefore reasonably expect that a quantum computer, built to simulate Nature, should be well-suited for these tasks.

- **Static properties.** Beyond dynamics, we can also ask about the behaviour of a quantum system in thermal equilibrium. This is typically phrased in terms of the Gibbs state, which is defined (for inverse temperature $\beta > 0$) by

  $$\rho_\beta := \frac{e^{-\beta H}}{\mathrm{tr}[e^{-\beta H}]} \ . \tag{1.6}$$

  As $\beta \to \infty$, this approaches the ground, i.e. lowest energy, state (if unique) of the Hamiltonian H, which we denote by $|\psi_0\rangle$. Properties of equilibrium properties are therefore given by expectation values of these states:

  $$\mathrm{tr}[O\rho_\beta] \ , \quad \langle\psi_0|O|\psi_0\rangle \ . \tag{1.7}$$

  A special case of particular interest is the ground state energy, $\lambda_0 = \langle\psi_0|H|\psi_0\rangle$. By definition, this is the lowest eigenvalue of H and describes the lowest energy that the system can possibly attain.

Static problems often arise as approximations for dynamical questions: in Nature, a small system of interest (such as a single molecule in chemistry) may exist in contact with an environment so large and chaotic that it cannot possibly be accounted for in exact dynamical calculations. The theorist can simplify this situation by assuming that their chosen subsystem will thermalise and arrive at a Gibbs or ground state in equilibrium. This assumption turns out to be extremely subtle, however: a priori, it is completely unclear exactly why and how these states should be produced as marginals via evolution of the form given in Eqs. (1.2)-(1.3). Recent developments in the theory of quantum Gibbs sampling have established conditions to prepare Gibbs states via local dissipative dynamics [33], however these require exponentially long times to arrive at such states for general Hamiltonians due to fundamental complexity-theoretic barriers. It is therefore not exactly clear the extent to which Nature is qualified to answer such questions on a reasonable timescale, and even less clear how well an artificial quantum computer should be expected to perform. This is a very important problem which we will discuss in detail in Section 2.

It is worth noting that the computation of static properties can often be algorithmically reduced back to dynamical tasks — for example, by attempting to prepare a Gibbs or ground state via adiabatic or dissipative methods [34, 35, 36, 33, 37], or by performing quantum phase estimation on a trial ground state [38].

Universal quantum computation

Beneath many layers of abstraction, all classical computation is fundamentally built on the manipulation of bits — simply two-state systems which can take 0 or 1 as their value. Many bits in parallel can be persuaded to solve interesting problems through the application of logical gates: transformations which alter the states of one or two bits according to rules such as AND, NOT, and OR. Classical circuits are constructed by applying many such gates in sequence to a set of bits; manipulation of information in this way provides a universal framework for classical computation.

Quantum computers work in much the same way, but with the word "quantum" inserted into all grammatically sanctioned whitespace. Rather than bits, we deal with quantum bits (familiarly known as "qubits"); these are physical systems with a two-dimensional Hilbert space $\mathcal{H}_1 \cong \mathbb{C}^2$, for which an orthonormal basis $|0\rangle$ and $|1\rangle$ is provided. A pure state $|\psi\rangle \in \mathcal{H}_1$ might be in either of these classical states, but also has the opportunity to occupy fundamentally quantum states between $|0\rangle$ and $|1\rangle$, of the form $|\psi\rangle = a|0\rangle + b|1\rangle$ for $a, b \in \mathbb{C}$ (normalised with $|a|^2 + |b|^2 = 1$). Rather than gates, quantum computers work by applying quantum gates. These similarly act on single qubits, or pairs of qubits, and are mathematically described as unitary transformations on the spaces $\mathbb{C}^2$ or $\mathbb{C}^2 \otimes \mathbb{C}^2$ respectively. Arbitrary quantum algorithms can then be executed on a system of n qubits $(\mathbb{C}^2)^{\otimes n}$ by the sequential application of quantum gates in what is correspondingly called a quantum circuit. At the end of the computation, a classical logical result can be extracted by performing a measurement. Since measurements in quantum mechanics are nondeterministic, this is a fundamentally probabilistic process and quantum circuits must be designed so that the desired computational outcome

(a) (b)

Figure 1.2: (a) A classical circuit taking two binary variables $x_1, x_2 \in \{0, 1\}$ as input, which inverts x_1 and then computes the logical AND of the two wires. (b) A quantum circuit on two qubits taking the state $|x_1\rangle \otimes |x_2\rangle$ as input, which executes a single-qubit gate on the first qubit followed by a controlled-NOT between the qubits. This circuit ends with probabilistic measurement of both qubits.

appears with sufficiently high probability to be detected. Some example classical and quantum circuits are illustrated in Fig. 1.2.

There are a couple of additional subtleties present in the quantum circuit model of computation which do not arise in the classical case. Firstly, the unitary group is continuous and hence infinite, which seems like it could pose a problem for the compact description of quantum circuits. In fact, this is not a major issue due to the Solovay-Kitaev theorem [39], which states that any local unitary operation can be approximated arbitrarily well from a (well-chosen) finite set of elementary gates.

The second issue is noise: any quantum device which is sufficiently precise to manipulate information on the level of qubits is necessarily sensitive enough to find itself subjected to all sorts of unwanted physical errors from the chaotic universe in which it sits. Protecting quantum information from noise is complicated by the no-cloning theorem, which prevents naive adaptations of classical error correction. Despite these barriers, the theoretical possibility of fault-tolerant quantum computation was ultimately demonstrated by the development of quantum error-correcting codes [40] and the subsequent landmark threshold theorem [41, 21]. Typically, quantum error correction leads to a very large overhead in terms of the physical resources required for computations, and the hunt for more efficient error correction protocols remains a central focus for theoretical research (see e.g. Ref. [42] for a recent review).

Digital quantum simulation

The liberal use of the word "quantum" leads to a structurally different and lucrative-sounding set of mathematical building blocks for computation, but it is perhaps not immediately obvious what the actual effect of this change should be at the algorithmic level: with an expanded set of local gates, what sort of new problems can we solve? As alluded to in the Section 1.1, one natural candidate is time-evolution of a quantum state under a local Hamiltonian. Solving this problem on a gate-based quantum computer is known as digital quantum simulation [43].

Given a n-qubit Hamiltonian H, the time-evolution operator e^{-itH} is highly non-local and does not typically correspond exactly to a quantum circuit built from local

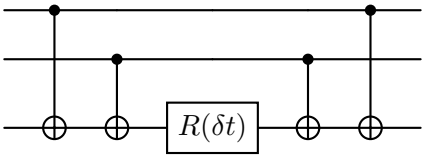


Figure 1.3: Quantum circuit implementing the 3-qubit unitary $e^{-i\delta t Z\otimes Z\otimes Z}$ for use within Trotter step for digital quantum simulation, where we define the single-qubit gate $R(\delta t) := \mathrm{diag}(1, e^{-i\delta t})$.

operations, but this can be overcome using a technique known as Trotterisation. This is the observation that, for a local Hamiltonian of the form $H = \sum_{A\subseteq[n]} H_A$, we can approximate

$$e^{-i\delta t H} \approx \prod_{A\subseteq[n]} e^{-i\delta t H_A} \ , \tag{1.8}$$

for very small time steps $\delta t \ll 1$. That is, we individually evolve under each term in the Hamiltonian one at a time. Each $e^{-i\delta t H_A}$ is a local operation, acting only on the sites $A \subseteq [n]$, and can be engineered by a local quantum circuit such as Fig. 1.3. If the H_A all commute then this approximation is exact, but otherwise it breaks down as δt increases. In order to simulate up to an arbitrary $t > 0$, we can simply divide the evolution up into N time steps of size $\delta t := t/N$, and use the Trotter approximation at each one, yielding

$$e^{-itH} \approx \left(\prod_{A\subseteq[n]} e^{-\frac{it}{N} H_A} \right)^N \ , \tag{1.9}$$

where the quality of the approximation improves as N increases. This is akin to how one might simulate physics on a classical computer: evolve in small time steps, updating each local component separately. In this way, the dynamics of an arbitrary k-local Hamiltonian can be approximately realised by a quantum circuit. Notice that the interaction graph of the circuit constructed in this way is identical to that of the simulated Hamiltonian, though this can be modified using additional gates overhead to shuttle qubits to the correct locations.

The following three decades of algorithmic development have yielded far more sophisticated algorithms for time evolution, with improved asymptotic circuit depth scalings [44, 45, 46, 47, 48], although the picture is somewhat subtle and asymptotically optimal algorithms are often not empirically favoured by resource estimates for concrete finite-sized problems of interest [30]. The focus of this thesis is not digital quantum simulation, so we will not discuss these aspects further: the important point is simply that efficient implementation of e^{-itH} is theoretically possible on a quantum computer, for a local Hamiltonian H.

Digital quantum simulation has several obvious benefits. Firstly, it is versatile: any local Hamiltonian can be simulated in this way from a universal gate set. Secondly, as

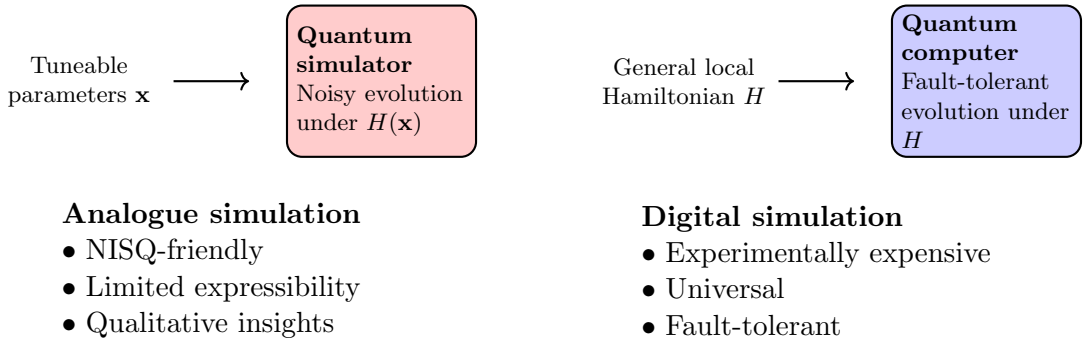


Figure 1.4: Comparison of the frameworks of analogue (left) and digital (right) quantum simulation.

mentioned, it can be made fault-tolerant: even in the presence of noise and experimental imperfections, one can protect a computation over arbitrarily long times via quantum error correction. The clear drawback to digital quantum simulation is the sheer difficulty of its experimental implementation. Resource estimates for concrete problems of interest in chemistry and physics [49, 50] suggest that millions of physical qubits may be required and controlled over timescales of days. Though hopefully achievable in the future, these specifications are well out of reach for present-day hardware.

Analogue quantum simulation

In light of the near-term difficulty of digital quantum simulation, analogue simulation has been proposed as an alternative with weaker experimental requirements, at the expense of a loss of universality and robustness [51]. We illustrate some differences between digital and quantum simulation in Fig. 1.4.

The basic idea of analogue simulation is as follows: we assume that the experimentalist has access to a tuneable device, capable of implementing a family of Hamiltonians $\{H'(\mathbf{x})\}$, where $\mathbf{x}$ is some list of tuneable parameters. For example, $H'(\mathbf{x})$ could be a Hamiltonian consisting of nearest-neighbour 2-local interactions in a 2D square lattice of qubits, where the interaction along each edge can be individually controlled, up to the physical constraints of the platform. We aim to simulate some target Hamiltonian H. If H can be realised exactly as an instance of $\{H'(\mathbf{x})\}$ then this is trivial; one simply chooses the correct parameters and allows the system to evolve naturally. However even when H cannot be directly identified as an element of $\{H'(\mathbf{x})\}$, it may be possible to choose $\mathbf{x}$ such that the behaviour of $H'(\mathbf{x})$ somehow mimics the desired properties of H — for instance, so that it has a provably similar ground state energy, or that its dynamics resemble those of H in some subspace. Using an analogue quantum simulator to study a system can be likened (and will be likened, in Section 3) to a civil engineer building a scale model of a project: this yields an object which is simpler to implement than the true system, but which approximately captures its essential qualitative properties.

The advantage of analogue quantum simulation is that its experimental requirements

are — at least in principle — much weaker than for the digital case. It is not necessary to construct a device capable of performing high-fidelity quantum gates from a universal set, but instead to evolve under the natural dynamics of the chosen Hamiltonian. There are a couple of obvious, yet rather subtle, drawbacks to this approach.

The first drawback is that of noise: since a Hamiltonian is directly implemented on the device, its dynamics do not occur in a logical space which is protected from errors. Inevitably, noise affecting the system will accumulate over time and lead to uncontrolled errors in observable expectation values. The propagation of such noise in analogue simulators has received theoretical attention in recent works, which have shown that its effects can be lessened slightly by stochastic cancellation of errors under some statistical assumptions on the noise [52, 53], or by restricting to the effects of noise within the past light cone of an observable of interest, rather than the whole system [54, 55]. Intuitively, one may expect that the effects of local noise on the system will not change the macroscopic behaviour of the simulation, but it is theoretically challenging to identify properties which are both robust to noise and classically difficult to simulate, and therefore quantum advantage for noisy analogue simulators remains a thorny mathematical issue. On the other hand, it can be heuristically argued that noise may actually be helpful for analogue quantum simulators; real world systems are themselves subject to unpredictable environmental noise, and thus a simulator with similar properties could still provide realistic insights. This statement is not rigorous, and in any case one must ensure that the noise on the simulator corresponds to a physically realistic noise model on the target Hamiltonian.

Secondly there is the problem of expressivity: given a highly constrained family $\{H'(\mathbf{x})\}$, what sorts of Hamiltonians H can we actually simulate? Clearly this depends strongly on how tunable one allows the simulator Hamiltonians to be, and exactly what is meant by the word "simulate". In a mathematically natural framework, Ref. [56] found that the (somewhat simplified) answer to this question is: surprisingly many! The fundamental tool for this purpose is perturbation theory. By choosing $\mathbf{x}$ carefully, the Hamiltonian $H'(\mathbf{x})$ may be engineered to have an invariant subspace (typically a low-energy space) upon which H arises as an effective theory. Such constructions were discovered in the context of Hamiltonian complexity theory [57], which provides an important (but less obvious) motivation to study the mathematics of analogue quantum simulation which we will address in much more detail in Section 2.

Despite these limitations, the experimental feasibility of analogue quantum simulation makes it attractive for near-term realisations, which have been demonstrated across a range of platforms at the scale of thousands of qubits (see, for example, Ref. [31] and references therein), well beyond the current capabilities of digital devices. Analogue simulators appear to be a key practical tool for the study and demonstration of quantum technology before the advent of large-scale fault-tolerant quantum computation.

2 Quantum physics — how hard can it be?

2.1 Classical and quantum complexity theory

In Section 1, we have introduced the basic problems of quantum simulation, and two distinct approaches (analogue and digital) to solving them on quantum devices. In this section we will discuss some concrete statements about the hardness of these problems through the framework of complexity theory, intended to motivate and compliment the subsequent parts of this thesis wherein complexity theory will not be explicitly discussed in much detail. After establishing some definitions, we will see in Section 2.2 that many problems in quantum simulation turn out to be rather trickier than we would like — that is, they are too difficult even for quantum computers in the worst case! In Sections 2.3 and 2.4, we will discuss whether this is truly a problem in physics, or merely an artefact of the unnecessary pessimism of worst-case complexity theory. Indeed, in Section 3 we will establish that certain types of complexity-theoretic reductions are unavoidably non-physical. Most of this section is a review of prior definitions and results, with the exception of the discussion surrounding Definition 2.4.1, which this Author has not seen explicitly stated elsewhere.

There is a lot of subtlety in the definitions of complexity classes which will be completely ignored in this thesis for the sake of conciseness. For a more rigorous treatment see for example Ref. [58], or Ref. [59] for a more quantum-focused review.

Decision problems and complexity classes

We begin by informally formalising the notion of a computational problem, specifically the framework of decision problems. A decision problem takes a finite string of bits $x \in \{0,1\}^*$ as its input, and as output one must decide whether x corresponds to a "yes" or "no" instance. A decision problem is therefore defined by disjoint sets $A_{\text{yes}}, A_{\text{no}} \subseteq \{0,1\}^*$. The union $A_{\text{yes}} \cup A_{\text{no}}$ might not comprise the entire space of possible bit strings $x \in \{0,1\}^*$; an algorithm is allowed to simply assume that $x \in A_{\text{yes}} \cup A_{\text{no}}$, and its output on the invalid x outside this set is unimportant. In this case the problem referred to as a promise problem, since the algorithm is "promised" that x will be a valid instance. The definition of x as a string of bits is a technicality; often more complicated objects are taken as input via some pre-agreed binary encoding. As an illustrative example, below we define the integer factorisation problem, whose classical hardness underpins the security of several cryptographic protocols.

Definition 2.1.1 (Integer factorisation)**.** Integer factorisation can be phrased as the following promise problem:

- **Input:** A pair $x = (x_1, x_2)$ of natural numbers $x_1, x_2 \in \mathbb{N}$.
- **Output:** Decide whether x_1 has a prime factor $p \in \mathbb{N}$ such that $2 \leq p \leq x_2$.

Formally speaking, the set A_{yes} consists of encodings $x = (x_1, x_2)$ such that x_1 has a prime factor smaller than x_2. A_{no} consists of the encodings $x = (x_1, x_2)$ such that this

is not the case. If x is not a valid encoding of two natural numbers, then it is neither in A_{yes} nor A_{no} as it does not satisfy the promise of the problem.

We categorise promise problems into complexity classes based on their tractability by various types of algorithms. Informally, a classical algorithm is defined as a family of classical circuits (one for every input length $|x|$), into which one loads the string x as input, and reads off the output (a zero or a one) from a specified output wire. If an algorithm outputs 1 for every input $x \in A_{\text{yes}}$, and outputs 0 for every input $x \in A_{\text{no}}$, then the algorithm is said to decide the promise problem $(A_{\text{yes}}, A_{\text{no}})$.

We say that an algorithm runs in time $T(|x|)$ if the depth (i.e., number of layers of logical gates) of this these circuits is upper bounded by $T(|x|)$ for all $|x|$, and that it runs in space $S(|x|)$ if the width (i.e., maximum number of wires simultaneously active) of these circuits is upper bounded by $S(|x|)$ for all $|x|$. In complexity theory, we typically consider an algorithm to be efficient if it runs in time scaling polynomially with $|x|$. We will write $\text{poly}(\cdot)$ as shorthand for any function which grows at most polynomially in its input.

Polynomial time P: A promise problem $(A_{\text{yes}}, A_{\text{no}})$ is in P if it can be decided by a classical algorithm which runs in time $\text{poly}(|x|)$.

Exponential time EXP: A promise problem $(A_{\text{yes}}, A_{\text{no}})$ is in EXP if it can be decided by a classical algorithm which runs in time $\exp(\text{poly}(|x|))$.

Polynomial space PSPACE: A promise problem $(A_{\text{yes}}, A_{\text{no}})$ is in PSPACE if it can be decided by a classical algorithm which runs in space $\text{poly}(|x|)$.

The class P encompasses all the problems which can be efficiently solved by deterministic classical computers. It is relatively straightforward to see that

$$\mathsf{P} \subseteq \mathsf{PSPACE} \subseteq \mathsf{EXP} \ . \tag{2.1}$$

To define the analogous quantum class to P, we must account for the randomness of quantum computation. Similarly to the classical case, a quantum algorithm is defined as a family of quantum circuits into which one loads the string x as input by preparing the computational basis state $|x\rangle$. Now, measuring the output wire will yield a non-deterministic result — this randomness is impossible to avoid, especially because one must restrict to a finite universal set of gates which may only approximately implement some desired set of operations. This is fixed by simply requiring the success probability to be above the constant $2/3$ (the specific number is completely arbitrary, and can be amplified closer to one by repeating the computation and taking a majority vote classically). We thus arrive at the complexity class BQP, which contains all problems admitting efficient quantum algorithms. In particular, the integer factorisation problem of Definition 2.1.1 lies in BQP due to Shor's algorithm [17].

Bounded-probability quantum polynomial time BQP: A promise problem $(A_{\text{yes}}, A_{\text{no}})$ is in BQP if there exists a quantum algorithm running in time $\text{poly}(|x|)$ such that the following holds. On inputs $x \in A_{\text{yes}}$ the output is 1 with probability $\geq 2/3$, and on outputs $x \in A_{\text{no}}$ the output is 0 with probability $\geq 2/3$.

There are a couple more complexity classes which will be useful for our purposes. These classes have an additional narrative structure known as an interactive proof system: for any input x, an all-powerful prover will produce a proof of the statement "$x \in A_{\text{yes}}$". This proof is given to an algorithm (the verifier) which must verify its correctness.

Classically, a proof takes the form of an additional bit string $v_x \in \{0,1\}^*$. For example, in the integer factorisation problem of Definition 2.1.1, v_x might be a binary representation of the factor p purported to divide x_1; it is difficult to find such a p, but easy to verify that a given p works. Problems whose proofs can be verified with polynomial-time classical algorithms define the class NP.

Non-deterministic polynomial time NP: A promise problem $(A_{\text{yes}}, A_{\text{no}})$ is in NP if there exists a classical algorithm on inputs (x, v_x) running in time $\text{poly}(|x|)$ such that the following holds. For any $x \in A_{\text{yes}}$, there exists $v_x \in \{0,1\}^*$ with $|v_x| = \text{poly}(|x|)$ such that the output of the algorithm on (x, v_x) is 1. Meanwhile for any $x \in A_{\text{no}}$, the output of the algorithm on (x, v_x) is 0 for all $v_x \in \{0,1\}^*$.

In the quantum case, the prover (who is referred to as Merlin due to their legendary proof-generation powers) gives the verifier (correspondingly Arthur, as in King Arthur) a quantum state $|\psi_x\rangle$ as a proof. Meanwhile Arthur is assumed, anachronistically, to have access to a quantum computer with which he can test the validity of this proof via a polynomial time quantum algorithm. This gives the class QMA.

Quantum Merlin Arthur QMA: A promise problem $(A_{\text{yes}}, A_{\text{no}})$ is in QMA if there exists a quantum algorithm taking inputs $(x, |\psi_x\rangle)$ running in time $\text{poly}(|x|)$ such that the following holds. For any $x \in A_{\text{yes}}$, there exists a state $|\psi_x\rangle$ on $\text{poly}(|x|)$ qubits such that the output of the algorithm on $(x, |\psi_x\rangle)$ is 1 with probability $\geq 2/3$. Meanwhile for any $x \in A_{\text{no}}$, the output of the algorithm on $(x, |\psi_x\rangle)$ is 0 with probability $\geq 2/3$ for all $|\psi_x\rangle$.

Initially, it may appear that the presence of an omnipotent sorcerer might make algorithm design redundant because they can simply be asked for the correct answer. The catch is thus: just as in real life, wizards are mischievous beasts who cannot always be trusted. For all instances $x \in A_{\text{yes}}$, Merlin will declare that the answer is yes and provide a valid proof of this fact. However, for invalid instances $x \in A_{\text{no}}$, the Wily Warlock will

still confidently proclaim that the answer is yes, and may provide a convincing-looking proof to attempt to fool Arthur into believing him. Thus Arthur is still faced with the highly non-trivial task of verifying the wisdom of the Grand Enchanter whilst detecting his deception; a more powerful Merlin can be both an asset and a problem for Arthur.

These are by no means the only complexity classes of interest to computer scientists; there are several variations on the Merlin-Arthur proof system whose relationship to QMA is not fully resolved (see Ref. [60] for a recent review), even extending to the case of multiple provers [61] and longer conversations between the prover and verifier [62, 63]. Arguing about the relative power of fictional wizards is a fascinating and fruitful area of research which has led to deep theoretical insights, from the PCP theorem [64] to the recent breakthrough result $\mathsf{MIP}^* = \mathsf{RE}$ [65].

Since quantum circuits can generally simulate classical circuits by simply restricting the gate set to classical operations, it is easy to see that

$$\mathsf{P} \subseteq \mathsf{BQP} \ , \quad \mathsf{NP} \subseteq \mathsf{QMA} \ . \tag{2.2}$$

Moreover, the existence of a prover can only increase computational power, since Arthur can always ignore whatever Merlin tells him and attempt to solve his problems on his own, implying that

$$\mathsf{P} \subseteq \mathsf{NP} \ , \quad \mathsf{BQP} \subseteq \mathsf{QMA} \ . \tag{2.3}$$

It is trivial to see that $\mathsf{BQP} \subseteq \mathsf{EXP}$, since any quantum computation can be simulated classically if one has enough time to write out and explicitly multiply the exponential-sized matrices and vectors describing the quantum gates and states. Less trivially, it can be shown that in fact BQP and QMA are both contained within PSPACE — this was initially proved as a corollary of the stronger result $\mathsf{QIP} \subseteq \mathsf{PSPACE}$ [62], though a more elementary proof is given by Ref. [66]. These inclusions, along with Eq. (2.1), are summarised in Fig. 2.1. It is widely suspected that all of these inclusions are strict, but only $\mathsf{P} \subsetneq \mathsf{EXP}$ is known, and it remains an open question whether even $\mathsf{P} \stackrel{?}{=} \mathsf{PSPACE}$.

Hardness and complete problems

Complexity classes are nebulous objects, defined abstractly as sets of all problems admitting some form of algorithmic solution. It is therefore not necessarily easy to get a good sense of what sort of problems might lie in a given class. Luckily, the study of complexity classes can often be reduced to the study of special promise problems, known as complete problems, which capture their overall behaviour.

We say (informally, here) that a promise problem $(A_{\text{yes}}, A_{\text{no}})$ is *hard* for a complexity class $\mathcal{C}$ if all problems in $\mathcal{C}$ can be reduced to instances of the hard problem in classical polynomial time. Intuitively, this means that $(A_{\text{yes}}, A_{\text{no}})$ is at least as difficult as every problem in $\mathcal{C}$; if it could be solved efficiently, then so could any problem in $\mathcal{C}$ by first reducing to the hard problem.

If a problem is hard for $\mathcal{C}$, and also contained within $\mathcal{C}$, then we say that it is *complete* for $\mathcal{C}$. Complete problems effectively capture the hardness of a complexity

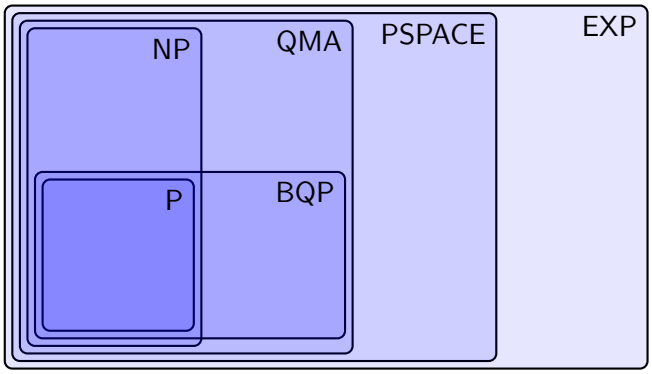


Figure 2.1: Known inclusion relations of the complexity classes defined in the text. Problems which can be efficiently solved on classical and quantum computers lie in the classes P and BQP respectively, occupying the bottom-left corner of the diagram. All of these inclusions are conjectured to be strict, but only $\mathsf{P} \subsetneq \mathsf{EXP}$ is known.

class, and serve as useful proxies for studying them. For example, proving that $\mathsf{P} \stackrel{?}{=} \mathsf{NP}$ is equivalent to proving the existence an efficient classical algorithm for any NP-complete problem — if such an algorithm existed, then any problem in NP could be reduced to the chosen NP-complete problem and then solved. Luckily, several complete problems are known for all of the above mentioned complexity classes. The canonical example of an NP-complete problem is k-Boolean satisfiability (k-SAT), described below.

Definition 2.1.2 (k-SAT)**.** For an integer $k \geq 1$, the promise problem k-SAT is defined by:

- **Input:** A Boolean formula on m variables $v_i \in \{0, 1\}$ of the form

$$x = C_1 \wedge C_2 \wedge \cdots \wedge C_n \ , \tag{2.4}$$

  where $\wedge$ denotes the logical "and", and each C_i is a Boolean clause containing at most 3 variables.

- **Output:** Decide whether the formula x can be satisfied for some assignment of the variables v_i, or whether x is unsatisfiable.

For $k \leq 2$, this problem can be decided in polynomial time but for $k \geq 3$ it is NP-complete, owing to the famous Cook-Levin theorem [67, 68], which is proved by reducing any classical computation to a Boolean formula such that the formula has some satisfying assignment if and only if the computation accepts some input. In the next section, we will see an analogous QMA-complete problem.

Theorem 2.1.3 (Cook-Levin theorem)**.** *k-SAT is NP-complete for $k \geq 3$.*

2.2 The hardness of simulation

Using the tools of Section 2, we are equipped to analyse some problems arising in quantum simulation through the lens of complexity theory. In this section we discuss some known containment and worst-case hardness results along these lines. See also Ref. [69] for an excellent, but not very recent, review.

Promise problems involving quantum simulation typically need to take a Hamiltonian as part of their input. It is therefore necessary that we restrict to a family of Hamiltonians which can be efficiently classically described. A convenient restriction is to k-local Hamiltonians, mentioned previously, which only contain terms acting on at most k spins, such that each term is normalised:

$$H = \sum_{A \subseteq [n]} h_A \quad k\text{-local} , \quad \|h_A\| \leq 1 . \tag{2.5}$$

There are only $\binom{n}{k} = \text{poly}(n)$ such subsets, so as long as each term h_A is specified to no more than $\text{poly}(n)$ bits of precision, the total Hamiltonian can be described with $\text{poly}(n)$ bits. The normalisation $\|h_A\| \leq 1$ is somewhat arbitrary and scaling h_A can be equivalently achieved by scaling the simulation time t, or the temperature for static problems.

Time evolution

Firstly, we can naturally phrase time dynamics as a promise problem, as given below.

Definition 2.2.1 (The k-local observable dynamics problem)**.** For $k \geq 1$, the k-local observable dynamics promise problem is defined by:

- **Input:** A k-local Hamiltonian H and a k-local observable O on n sites, both of the form given in Eq. (2.5), a real simulation time $t = \text{poly}(n)$, and real numbers α, β such that $\beta - \alpha \geq 1/\text{poly}(n)$.
- **Output:** Decide whether the quantity

$$\langle O \rangle(t) = \langle 0^{\otimes n} | e^{itH} O e^{-itH} | 0^{\otimes n} \rangle \tag{2.6}$$

  satisfies $\langle O \rangle(t) \leq \alpha$ or $\langle O \rangle(t) \geq \beta$, under the promise that one of these is the case.

This problem is very general, and captures all questions of the form "after time t, what will the value of ... be?" It is perhaps not surprising, then, that this problem turns out to be complete for universal quantum computation.

Theorem 2.2.2. *The k-local observable dynamics problem is* BQP*-complete for any constant $k \geq 2$.*

The fact that k-local observable dynamics are contained within BQP follows immediately from the discussion of the discussion of Section 1.2; by digital quantum simulation,

the state $e^{-itH}|0^{\otimes n}\rangle$ can be prepared in polynomial time on a quantum computer, and the hence the value of $\langle O\rangle(t)$ can be estimated to $1/\operatorname{poly}(n)$ precision by repeated measurements.

The hardness of this problem is somewhat more subtle: one must show that arbitrary discrete quantum computations can be encoded into continuous-time dynamics under a local Hamiltonian. The key idea for this result is very old, in fact due to Feynman [70], before the notions of quantum complexity theory had been formalised.

By definition, any problem in BQP can be solved, for an input x of length $|x| = n$, by applying a quantum circuit on $N = \operatorname{poly}(n)$ qubits consisting of $L = \operatorname{poly}(n)$ 2-qubit gates to an initial state $|0^{\otimes N}\rangle$, and performing a measurement on a single output qubit. We can represent the quantum circuit as a sequence of gates $U_L U_{L-1} \dots U_2 U_1$, and define the Hamiltonian H on a system of $L+1$ clock qubits (labelled $0, 1, \dots, L$) and N logical qubits by

$$H = \sum_{l=1}^{L} \left(T_l \otimes U_l + T_l^\dagger \otimes U_l^\dagger \right) \tag{2.7}$$

where T_l acts as the operator $|110\rangle\langle 100|$ on the clock qubits $l-1, l, l+1$, and the special case of T_L acts as $|11\rangle\langle 10|$ on clock qubits $L-1, L$. The key observation is that H preserves the subspace spanned by states of the form

$$|\psi_l\rangle = |1^{\otimes(l+1)} 0^{\otimes(L-l)}\rangle \otimes \left(U_l U_{l-1} \dots U_2 U_1 |0^{\otimes N}\rangle \right) \ , \quad l = 0, 1, 2, \dots, L \ . \tag{2.8}$$

The first $L+1$ qubits keep track of what stage in the computation the N-qubit logical system is in, whilst the remaining N qubits implement the circuit. The only transitions allowed by the Hamiltonian are those which advance the clock by one step, and apply the corresponding unitary to the logical qubits. That is, transitions of the form $|\psi_l\rangle \leftrightarrow |\psi_{l+1}\rangle$. In the subspace spanned by the $|\psi_l\rangle$, H acts as the matrix

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & 1 & 0 & & 0 & 0 \\ 0 & 1 & 0 & 1 & & 0 & 0 \\ 0 & 0 & 1 & 0 & & 0 & 0 \\ \vdots & & & & \ddots & & \vdots \\ 0 & 0 & 0 & 0 & & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 \end{pmatrix} \ . \tag{2.9}$$

Since we initialise in the state $|\psi_0\rangle = |0^{\otimes(L+N)}\rangle$, time evolution under H results in transitions through the states $|\psi_l\rangle$, and it can be shown that after $\operatorname{poly}(n)$ time the probability of measuring the Lth clock bit in its "1" state (and hence the N logical bits in the correct post-computation state) is at least $1/\operatorname{poly}(n)$. In fact, this probability can be boosted close to one by tuning the coefficients of H using the theory of perfect state transfer [71].

Note that H is 5-local, and thus does not quite suffice for the $k \geq 2$ result advertised in Theorem 2.2.2. This can be obtained by perturbative simulation approaches which we discuss later in this section, at the expense of introducing polynomially scaling local coefficients in the Hamiltonian. Geometric locality in 2D can also be obtained by adding additional gates to the circuit and with suitable placement of the clock qubits as shown in Ref. [72], which also establishes strict planarity through perturbation theory.

The problems of 2-locality and geometric constraints are both solved by more recent, and arguably more elegant, constructions based on quantum walks [73, 74]. These works construct planar 2-local Hamiltonians with only one type of interaction (hence defined only by their interaction graph) which can simulate arbitrary quantum circuits. The intuitive idea is that the quantum circuit is laid out in physical space, with logical quantum information dynamically propagating through. In fact, using different ideas with cellular automata, BQP-complete computations can be encoded even into the dynamics of translationally-invariant Hamiltonians on a line [75, 76], albeit with enlarged local dimension.

The BQP-completeness of Hamiltonian time evolution can be viewed as fairly convincing evidence for exponential quantum advantage for quantum simulation tasks; if any polynomial classical algorithm existed, this would imply that all quantum computation could be efficiently simulated, i.e. $\mathsf{BQP} \stackrel{?}{=} \mathsf{P}$. This would be a very surprising, not to mention disappointing, collapse of complexity classes.

Of course, for the classical difficulty of simulation to manifest itself, it is necessary that truly quantum effects are present in the dynamics. For systems which remain only weakly entangled over the course of the simulation, classical tensor network methods perform very well. We won't discuss classical algorithms in detail here, but just make the following qualitative remarks: in 1D, tensor network methods for time evolution work very (in fact provably) well for weakly entangled systems [77, 78]. In higher dimensions, this picture is a little more complicated: exact contraction of tensor networks becomes a computationally intractable problem (even for quantum computers!) [79], so one must resort to heuristics [80, 81].

Ground states and the local Hamiltonian problem

We now turn to the setting of static properties, and focus on ground states. For a local Hamiltonian H, we denote its eigenvalues by $\lambda_0(H) \leq \lambda_1(H) \leq \lambda_2(H) \leq \ldots$. Given a system defined by a local Hamiltonian H and coupled to a low-temperature bath, it is natural to expect that over enough time it will relax to a ground state $|\psi_0\rangle$ (especially when there is a spectral gap), and hence settle with energy $\lambda_0(H)$. Finding $\lambda_0(H)$ is the famous local Hamiltonian problem, phrased as a promise problem below.

Definition 2.2.3 (k-local Hamiltonian problem)**.** For $k \geq 1$, the k-local Hamiltonian problem is defined by:

- **Input:** A k-local Hamiltonian H as in Eq. (2.5), and real numbers α, β such that $\beta - \alpha \geq 1/\operatorname{poly}(n)$.

- **Output:** Decide whether the ground state energy $\lambda_0(H)$ satisfies $\lambda_0(H) \leq \alpha$ or $\lambda_0(H) \geq \beta$, under the promise that one of these is the case.

Note that, in contrast to Definition 2.2.1, we restrict to measuring the observable H rather than arbitrary ground state properties $\langle\psi_0|O|\psi_0\rangle$. This is mostly for a practical reason: the ground state of H may not be unique, leaving the latter problem ill-defined. With an additional promise that H has a unique ground state below with energy below β when one exists, it is straightforward to see that these problems are equivalent.

Notice that 3-SAT can be realised as a special case of the 3-local Hamiltonian problem: any Boolean formula $C_1 \wedge C_2 \wedge \cdots \wedge C_n$ on m variables can be viewed as a local Hamiltonian on m qubits, by associating each variable to a qubit, and for each C_i including a term which projects onto the non-satisfying assignments of the variables included in C_i. If the Boolean formula has a satisfying assignment, then the corresponding computational basis state will have energy $\lambda_0 = 0$, whilst if no such assignment exists then the Hamiltonian will have $\lambda_0 \geq 1$, since at least one clause will be unsatisfied by any computational basis state. This establishes the NP-hardness of the k-local Hamiltonian problem for $k \geq 3$. In fact the analogy to k-SAT goes deeper than this; the k-local Hamiltonian problem turns out to be QMA-complete for any $k \geq 5$ due to a quantum version of Theorem 2.1.3 initially due to Kitaev [39], which was extended to $k \geq 2$ by subsequent works [82, 57]:

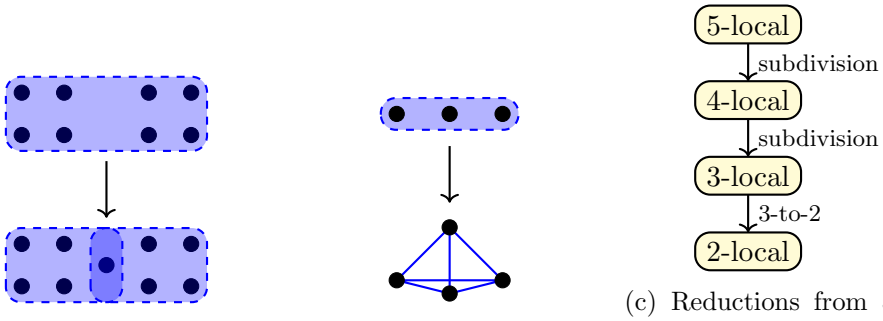
Theorem 2.2.4 (Quantum Cook-Levin theorem)**.** *The k-local Hamiltonian problem is* QMA*-complete for any constant $k \geq 2$.*

The containment of the k-local hamiltonian problem in QMA is relatively straightforward: Merlin can simply give Arthur the ground state as a proof, which Arthur can measure to verify that its energy is below α. Meanwhile, the QMA-hardness follows for $k = 5$ follows from a modification (due to Kitaev [39]) of the clock Hamiltonian introduced in Eq. (2.7); we will skip the details here; see for example Ref. [83], Section 1.3, for a description of this construction and some variants. In particular, for any circuit $U_L U_{L-1} \ldots U_2 U_1$, this procedure yields a 5-local Hamiltonian whose ground state is zero if it has an accepting input (one which outputs $|1\rangle$ on a specified wire), and at least $\sim L^{-3}$ otherwise.

Perturbative reductions of local Hamiltonian problems

In the previous section, we stated Theorem 2.2.4 as a hardness result for 2-local Hamiltonians, but slyly only sketched the proof in $k = 5$ — indeed we pulled a similar trick for Theorem 2.2.2. In fact, the construction for $k = 5$ can be reduced to $k = 2$ with one conceptual trick: perturbation theory. We sketch this approach here for completeness, but omit most technical details as these will be examined in Sections 3 and 4.

The basic idea is to take a hard instance of the 5-local Hamiltonian problem, and locally substitute each 5-local term with a Hamiltonian known as a gadget. Gadgets are constructed by introducing additional ancillary qubits to the system, which are energetically restricted to their $|0\rangle$ space by the presence of very strong 1-local terms of



(a) Subdivision gadget. (b) 3-to-2 gadget. (c) Reductions from 5-local to 2-local.

Figure 2.2: (a) Interaction hypergraph before and after application of the subdivision gadget to a 8-local interaction. This gives a gadget Hamiltonian with locality $\lceil 8/2 \rceil + 1 = 5$. (b) Interaction hypergraph before and after application of the 3-to-2 gadget to a 3-local interaction. We draw 2-local terms are drawn as lines for clarity. (c) Gadget reductions from a general 5-local Hamiltonian to a 2-local Hamiltonian.

the form $|1\rangle\langle 1|$. By including (weaker) interactions between these ancillary qubits and the original system, they can mediate effective interactions in the low-energy space of the overall system. The particularly relevant case for us here is the fact that gadgets containing 2-local interactions can give rise to an effective Hamiltonian containing 5-local interactions. Thus we can reduce the 5-local Feynman-Kitaev history state Hamiltonian to a 2-local gadget Hamiltonian which (up to a small error controlled by the perturbative approximation) has the same ground state energy.

Such techniques were originally used by Ref. [57], however we describe the approach of the slightly later work Ref. [72]. There the Authors introduce several perturbative gadget constructions, including two specifically for the task of locality reduction: the subdivision gadget and the 3-to-2 gadget, illustrated in Fig. 2.2. Subdivision gadgets can be used to reduce any k-local Hamiltonian to a $(\lceil k/2 \rceil + 1)$-local Hamiltonian; applying this construction twice allows a reduction from 5-local to 3-local. 3-to-2 gadgets can then be used to reduce 3-local interactions to 2-local gadget Hamiltonians, as required. More detail about these gadgets appears in Section 3.4.2.

In fact, the power of perturbative gadgets can be taken much further than locality reduction. In Ref. [72], the authors also use gadgets to simplify the interaction graph of the Hamiltonian to a strictly planar geometry by implementing gadgets which effectively couple non-adjacent sites in a lattice. Later, Refs. [84, 85, 86] introduced further gadgets allowing stronger statements than Theorem 2.2.4, restricting H to only contain certain types of interaction such as Heisenberg-type $X \otimes X + Y \otimes Y + Z \otimes Z$. A full complexity-theoretic classification of local Hamiltonian problems was established in Refs. [86, 87], showing that restricting to a given set of interactions leads to either QMA-completeness, or one of three other levels of hardness. Through different techniques, QMA-completeness has been established even for one-dimensional systems [88, 89], and even in the case of translational invariance [90]. These require enlarging the local Hilbert space dimension;

for example Ref. [89] establishes QMA-completeness for 2-local Hamiltonians on a line of 8-dimensional qudits. By interpreting each qudit as a system of 3 qubits, one can view such a system as a 6-local qubit Hamiltonian which can be perturbatively reduced to a 2-local system at the expense of strict 1D connectivity.

Universal quantum simulation

The techniques described above allow one Hamiltonian to reproduce the ground state of another by perturbatively engineering the desired interactions as an effective low-energy theory. As observed in Ref. [87], and extended by Ref. [56], this analysis can be extended much further than the ground state alone: in fact the entire spectrum of the original Hamiltonian can be realised as the low-energy spectrum of the perturbative gadget Hamiltonian.

This gives a rigorous foundation to study analogue quantum simulation from the theoretical perspective: given a Hamiltonian H' which reproduces the spectrum of a different Hamiltonian H as a low-energy theory, H' can be said to simulate H. In Ref. [56], this concept was taken slightly further to include the notion of locality of simulation: informally this gives a requirement that local interactions in H are simulated locally on H' (see Section 3.2 for technical details). This locality condition is satisfied automatically for gadget-type Hamiltonians, which are modular by construction, and has the convenient consequence that any local noise model on the simulator H' can be equivalently viewed as local noise acting on H.

In Ref. [56], this definition is applied to show that very simple families of Hamiltonians are universal, in the sense that they can simulate all other many-body Hamiltonians through multiple rounds of perturbation theory analogous to the local Hamiltonian problem reductions above. The fact that the entire spectrum is simulated has many nice consequences: ground states, Gibbs states, and time evolution are all meaningfully simulated by the gadget Hamiltonian.

The disadvantage of this approach, which we will expand on in Section 2.4, is that it leads to pathological Hamiltonians which are impossible to realise in practice. The problem is interaction strengths: each round of perturbation theory requires interactions which scale polynomially in the system size (and inverse-polynomially in the desired accuracy) in order to achieve the necessary energy gap for perturbation theory to converge, and over multiple rounds the degree of these polynomials becomes very high. In Section 3 we will see that these strong interactions are quite generally unavoidable for Hamiltonian simulations which reduce locality, while in Section 4 we will show how non-criticality can be exploited to reduce these costs with classical postprocessing techniques.

In Ref. [91], it was proved that the QMA-completeness of a family of Hamiltonians is equivalent to its universality as a class of simulators, and therefore these two notions of expressivity are in correspondence. One might wonder whether these expressivity conditions are also equivalent to the BQP-completeness of the associated time-dynamics problem (as in Definition 2.2.1), but this is not the case: there exist families of Hamiltonians with BQP-complete dynamics (via quantum walks) [74] but whose ground state energy problem is complete for StoqMA [92], a complexity class which is expected to be

smaller than QMA [93].

Thermal states

Having discussed the hardness of ground state energy problems, it is natural to extend this discussion to thermal states, which will be particularly relevant in Section 4. In some sense, thermal states appear simpler to deal with: for a Hamiltonian H and inverse temperature $\beta \geq 0$, the Gibbs state $\rho_\beta := e^{-\beta H}/\operatorname{tr}[e^{-\beta H}]$ is always a well-defined and continuous function of H, in contrast to ground states which are not necessarily unique. A natural computational problem for Gibbs states (which can be viewed as an imaginary-time analogue version of Definition 2.2.1) is the computation of a k-local observable O on ρ_β. We define this explicitly below.

Definition 2.2.5 ((k,β)-Gibbs state expectation problem)**.** For $k \geq 1$ and a non-decreasing function $\beta(n) \geq 0$, the (k,β)-Gibbs state expectation problem is defined by:

- **Input:** A k-local Hamiltonian H and a k-local observable O, both of the form given in Eq. (2.5), and real numbers α, β such that $\beta - \alpha \geq 1/\operatorname{poly}(n)$.
- **Output:** Decide whether the Gibbs state expectation value, defined as

$$\langle O \rangle := \frac{\operatorname{tr}[e^{-\beta H} O]}{\operatorname{tr}[e^{-\beta H}]} \ , \tag{2.10}$$

  satisfies $\langle O \rangle \leq \alpha$ or $\langle O \rangle \geq \beta$, under the promise that one of these is the case.

The computational complexity of this problem turns out to depend strongly on the value of inverse temperature β. For β very small, the Gibbs state ρ_β approximates the maximally mixed state, $\rho_\beta \approx \tau \propto \mathbb{1}$, for which observable estimation is trivial. It turns out that (assuming H has bounded interaction degree) this easiness regime extends up to a constant positive threshold $\beta_* > 0$ [94, 95], below which the state ρ_β remains unentangled and easy to classically simulate. The energy of ρ_β approximates the ground state energy up to a correction of order $\sim n\beta^{-1}$, and thus for $\beta \sim \operatorname{poly}(n)$ and $k \geq 2$ the (k,β)-Gibbs state expectation problem is at least QMA-hard by reduction from the k-local Hamiltonian problem. As we will mention in Section 2.4, this hardness would extend to the regime of constant $\beta > 0$ if the quantum PCP conjecture holds. To the Author's knowledge, it is not known whether this problem is complete for QMA in the low-temperature regime, or whether it may be even harder.

The QMA-hardness of this problem immediately implies (assuming $\mathsf{BQP} \overset{?}{\neq} \mathsf{QMA}$) that Gibbs states like ρ_β cannot, in general, be efficiently prepared on quantum computers. Recent breakthroughs [37, 33] have established the existence of quantum Gibbs samplers — local dissipative processes, which can be simulated on quantum computers and which provably converge to the Gibbs state of a given Hamiltonian — however these require exponential time to converge in general.

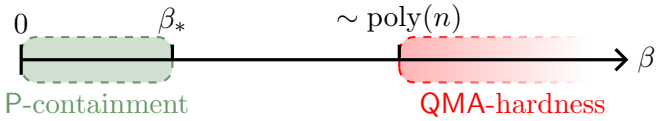


Figure 2.3: Hardness of the (k, β)-Gibbs state expectation problem, for different values of the inverse temperature $\beta \geq 0$. The high-temperature case ($\beta \leq \beta_*$, for constant $\beta_* > 0$) is classically easy [94, 95], whereas the low-temperature case ($\beta \sim \text{poly}(n)$, and possibly at a constant threshold if the quantum PCP conjecture holds [96]) is QMA-hard by reduction to the local Hamiltonian problem.

The electronic structure problem

So far, we have dealt with rather abstract Hamiltonians defined on systems of qubits — here we mention a problem of more obvious practical importance for chemistry.

Typically, chemists are interested in the dynamics of molecular systems containing nuclei and electrons in order to find stable configurations of molecules and dynamical pathways for chemical reactions to occur. The overall energy for a molecular system depends on the coordinates $\mathbf{r}$ of the electrons and coordinates $\mathbf{R}$ of the nuclei, leading to a Hamiltonian of the form

$$H(\mathbf{r}, \mathbf{R}) = H_{\mathsf{el}}(\mathbf{r}) + H_{\mathsf{nuc}}(\mathbf{R}) + H_{\mathsf{el,nuc}}(\mathbf{r}, \mathbf{R}) \ , \tag{2.11}$$

where $H_{\mathsf{el}}(\mathbf{r})$ and H_{nuc} contain the kinetic energy and Coulomb interactions of the electrons and nuclei respectively, whilst $H_{\mathsf{el,nuc}}(\mathbf{r}, \mathbf{R})$ contains the interactions between electrons and nuclei. This problem is simplified via the Born-Oppenheimer approximation, which takes advantage of the fact that electrons are much less massive than nuclei and therefore can be expected to settle into an equilibrium state very quickly whilst the nuclei behave essentially classically. In fact, once the coordinates $\mathbf{R}$ of the nuclei are fixed, we assume that the electrons whizz around with such zeal that they settle into their lowest-energy configuration with respect to the external nuclear potential before the nuclei have time to react. This means that — assuming that the electronic ground state energy can be calculated for every configuration of nuclei $\mathbf{R}$ — the motion of the nuclei can be approximated through classical equations of motion under a Hamiltonian depending only on $\mathbf{R}$.

Of course, "assuming that the electronic ground state energy can be calculated for every configuration of nuclei" is a rather strong statement; as we have seen in the preceding discussion, ground state problems tend to be quite difficult. The problem of computing electronic ground state energies is known as the electronic structure problem. We will not discuss the details of this problem, which are not relevant to this thesis, but highlight a couple of important points:

- The electronic structure problem involves minimisation over an infinite-dimensional Hilbert space, since the coordinates $\mathbf{r}$ exist in continuous space, and must be discretised through a procedure known as second quantisation. This step is typically

ignored for the purposes of complexity theory, and a finite-dimensional second-quantised Hamiltonian is given as input to the electronic structure problem [85, 97].

- The electronic structure problem is distinct from the local Hamiltonian problems discussed above, as electrons obey fermionic statistics: instead of a system of qubits, the Hilbert space on which the Hamiltonian acts is a Fock space generated by creation and annihilation operators. There are well-established methods for transforming fermionic systems to qubit systems [98, 99, 100]; typically these incur some additional overhead for digital quantum simulations, and can be optimised for specific problems of interest [101, 102].

Alas, the electronic structure problem is also known to be QMA-complete. This was initially shown by Ref. [85], by reduction from Fermi-Hubbard Hamiltonians with an external magnetic field. The alternative approach of Ref. [97] avoids the external magnetic field, and instead uses the choice of electronic orbitals to encode the computational problem. In both cases, it is important to note that QMA-hard instances are constructed via a series of perturbative reductions which result in highly artificial Hamiltonians — the question of whether these QMA-hardness results accurately reflect the difficulty of real physical systems will be discussed in Section 2.3.

Although the electronic structure problem is well-motivated and extremely useful from the perspective of quantum chemistry, we should note the following very simple observation. Before the Born-Oppenheimer approximation was applied, the joint nucleus-electron dynamics problem was in principle within BQP and amenable to digital quantum simulation, but somehow our attempt at a simplifying approximation elevated its hardness to QMA. When working with quantum computers, it may seem more natural to directly evolve in time under the Hamiltonian of Eq. (2.11). Indeed this has been studied [103, 104] and in principle allows the simulation of chemical dynamics in polynomial time — however the additional nuclear degrees of freedom lead to fairly expensive resource requirements [105].

2.3 Is physics worst-case hard?

In Section 1, we motivated the problem of quantum simulation as a key target for quantum computers to obtain computational speedups over their classical counterparts. Section 2.2 appears to suggest that these problems are rather harder than we had bargained for (QMA-hard), and in fact may not be tractable even with a hypothetical quantum machine! In the following sections we discuss some reasons to believe that this may not be the case, and the consequences for quantum simulation.

The computational limits of Nature

Purely on physical grounds, we may suspect that it is unreasonable for QMA-complete problems to appear in the real world: in Section 1.2 we saw that local Hamiltonian dynamics can be simulated on a quantum computer. Therefore, such dynamics cannot solve QMA-hard problems in polynomial time unless $\mathsf{BQP} \stackrel{?}{=} \mathsf{QMA}$.

A particularly striking perspective on this fact comes from considering a complexity class not mentioned above: QCMA. This has the same definition as QMA, except Merlin's proof comes in the form of a classical bit string, rather than a quantum state [106]. In particular, QCMA includes any local Hamiltonian problem where the ground state (or a procedure to prepare it) can be described classically in polynomial space. It is clear that QCMA $\subseteq$ QMA, and it is suspected that this inclusion is strict [60, 107]. Assuming this is the case, QMA-hard Hamiltonians have ground states which cannot be prepared by any polynomial-sized quantum circuit (which Merlin could otherwise use as a classical proof). This means that, even if Nature knew a ground state of a hard instance H in advance and actively *tried* to arrive at it, this would be impossible in polynomial time without violating the Schrödinger equation Eq. (1.2) (or more generally, dissipative dynamics under Lindblad's equation [108, 109], which we do not discuss here).

Based on this, it seems unlikely that solving QMA-hard problems is necessary to understand real physical systems. On the other hand, ground state energy are ubiquitous in computational physics and chemistry and appear to yield excellent predictions in practice. There are two possible resolutions to this tension:

1. Nature is not actually finding the ground state. This seems like a reasonable concession to make for real systems which exist at positive temperature, and is perhaps a physically necessary concession in light of the above discussion. Often, for instance in the electronic structure problem, QMA-hard problems appear as "simplifications" of dynamic problems which are known to be in BQP. Of course, this is a genuine simplification for small problems of interest where the complexity of the dynamical simulation outweighs the asymptotic hardness of the ground state energy, and is particularly understandable in the context of classical simulation where dynamics is already hard. When analysing asymptotic complexity, however, it might be misleading to reduce to a problem which is considerably harder in the worst case.

2. The Hamiltonians appearing in Nature do not have QMA-hard ground state energy problems. This conclusion seems to have some empirical support from decades of computational chemistry, as we discuss below. This possibility is quite intriguing, as it suggests that there might be some additional structure present in the sorts of problems that appear in the real world which precludes worst-case hardness. Mathematically classifying this structure, and how it can be exploited to facilitate efficient computation, is an important problem to understand the power of both classical and quantum approaches to practical quantum simulation. We discuss some ideas along these lines in Section 2.4.

The surprising power of classical computation

In Point 2 above, we mentioned that many practical problems appearing in physics and chemistry might not be QMA-hard. The optimistic conclusion to draw from this might be that many such problems are likely to be easy for quantum computers, facilitating exponential speedups over classical attempts at the same tasks. However this statement

hides an important assumption: for exponential quantum advantage, we require problems to be classically hard! If we reduce to a subclass of problems contained within BQP, we must consider the possibility that they are also contained within P, and hence classically tractable.

Indeed, motivated by a conspicuous lack of quantum computers over the last half-century, computational chemists have devised sophisticated classical heuristics for ground-state problems which often turn out to be very effective in practice. A recent specific example of this includes Ref. [110], which made substantial progress towards classically solving the FeMo-cofactor model, a long-standing benchmarks for quantum algorithms [111, 112] due to assumptions of its classical intractability. Particularly for ground state problems in chemistry, there is some debate about whether we should actually expect exponential quantum advantage to be exhibited in practice; for example Ref. [112] provides numerical evidence that many systems which appear amenable to quantum computers also admit good classical heuristics.

It is also worth noting that there is an asymmetry in the comparison of Ref. [112]; on the classical side, a suite of heuristics are used, developed over decades of quantum chemistry research and adapted based on empirical data to fit specific systems, whilst on the quantum side only simple approaches with rigorous guarantees (quantum phase estimation on a state prepared by a classical heuristic, and adiabatic state estimation) are considered. One might wildly speculate that the availability of large-scale quantum computers might facilitate the development of heuristic quantum algorithms, which offer vast (polynomial, if not exponential) improvements over the best known classical methods for practical instances of NP-hard and QMA-hard problems despite a lack of theoretical guarantees, however such a claim can only be validated empirically.

Complexity theory and complexity practice

Of course, statements of the form "Nature is QMA/BQP/NP-hard" are nonsensical: the real world is an impossibly large physical system which typically does not pose explicit and clean computational problems with interesting asymptotic limits. On the contrary, practical systems of interest come at fixed sizes, for which concrete resource estimates are more important than complexity classifications: QMA-hardness isn't an obstruction if we're only interested in the case $n = 5$, and P-containment isn't useful if the only known algorithm takes time $\sim n^{20}$.

So why bother studying the complexity of quantum simulation at all? Firstly, it's the only real way to make sense of things from a mathematical perspective, and provides rigorous evidence for the utility of quantum computers without having to build them first. Secondly, and more importantly, studying complexity theory can give far deeper insights into problems than merely classifying their hardness. By constructing worst-case instances (and identifying easier instances) we can examine them to determine the features that make them hard in order to understand where the fundamental limits of algorithm design lie, and what structure must be exploited by heuristic methods.

2.4 Optimism in local Hamiltonian problems

In light of the discussion of Section 2.3, in this section we mention some physically-motivated variants of the local Hamiltonian problem which may avoid worst-case hardness. We will revisit some of these assumptions in later parts of the thesis, however this discussion will raise several open questions which will regrettably remain unanswered. Some key results are summarised in Figure 2.4.

Guiding states

We have seen that the local Hamiltonian problem can be solved with access to an all-powerful wizard who can provide the ground state $|\psi_0\rangle$ with perfect precision. Of course, this ability of Merlin's is highly non-trivial (the dimension of a system of n qubits is 2^n, making exhaustive search completely infeasible), though for many practical problems of interest — in chemistry for example — classical methods empirically seem to give fairly good approximations. We can therefore imagine a situation in which Merlin is replaced by a skilful chemist who supplies us with instructions to prepare a state $|\phi\rangle$ which, though not necessarily the exact ground state, is a useful hint for a quantum computer. The state $|\phi\rangle$ is accordingly referred to as a guiding state.

This leads to the guided local Hamiltonian problem [113]: this is a variant on the local Hamiltonian problem of Definition 2.2.3 where the input includes instructions to prepare a guiding state $|\phi\rangle$, which is promised to have some overlap $\eta := |\langle\phi|\psi_0\rangle| > 0$ with a ground state $|\psi_0\rangle$ of H. Typically η is assumed to be at least $1/\operatorname{poly}(n)$ (a random guess would yield $e^{-\Omega(n)}$), which is sufficient to solve the ground state energy problem on a quantum computer via quantum phase estimation [114]. Quantum phase estimation provides a method to take a general state $|\phi\rangle$, expanded in the eigenbasis $\{|\psi_j\rangle\}_{j\geq 0}$ of H,

$$|\phi\rangle = \sum_{j\geq 0} c_j |\psi_j\rangle \ , \tag{2.12}$$

and transform it to a state of the form

$$|\phi\rangle \mapsto \sum_{j\geq 0} c_j |\psi_j\rangle \otimes |\hat{\lambda}_j\rangle \ , \tag{2.13}$$

where $|\hat{\lambda}_j\rangle$ is a state on a second register corresponding to a classical binary expansion of the eigenvalue λ_j of H. Thus the values of $\hat{\lambda}_j$ can be sampled by measuring the second register, each with probability $|c_j|^2$. Hence, in time $\sim 1/|c_0|^2 \sim \eta^{-2}$, the ground state energy can be obtained.

The complexity of the guided local Hamiltonian problem was introduced by Ref. [113], and subsequent work [115] established that it is BQP-complete even for very large overlap $\eta \sim 1 - 1/\operatorname{poly}(n)$, and for 2-local Hamiltonians in 2D. This result tells us exactly where we can expect quantum computers to offer solutions to local Hamiltonian problems: whenever good guiding states can be guessed and prepared on the quantum device. Sadly,

the abilities of mortal chemists are widely expected to fall short of those of omnipotent wizards; assuming $\mathsf{QMA} \stackrel{?}{\neq} \mathsf{BQP}$, many Hamiltonians do not have efficiently guessable guiding states. Moreover, situations where a good guiding state ansatz can be guessed often coincide with effective classical heuristic methods, leading to a somewhat fuzzy picture from the perspective of trying to identify regimes for quantum advantage [112].

Non-critical systems

One factor which often appears to be present in hard instances of ground state problems is criticality. For ground state problems, criticality can be characterised by the closing of the spectral gap of a family of Hamiltonians as the system size increases. Non-critical systems, in which the spectral gap remains lower bounded by a constant as the system size grows, are known to exhibit exponentially decaying correlations [116]. This allows the reduction of global ground state energy problems to the solution of many local constraints [117], and in one spatial dimension it is even possible to solve local gapped Hamiltonians in polynomial time classically [118].

Beyond 1D, the consequences of non-criticality for complexity theory are not fully understood. In Ref. [119], it was shown that history state Hamiltonians (i.e., those which naturally encode a QMA verification circuit, for some definition of "naturally") are generally critical — thus a completely new idea is necessary if QMA-hardness is to be shown for non-critical Hamiltonians. Moreover, for a modification of the local Hamiltonian problem where the promise gap is taken to be exponentially small (that is, $\beta - \alpha \sim e^{-\Omega(n)}$ in the notation of Definition 2.2.3), the promise of an inverse-polynomial spectral gap results in reduced hardness of the problem [120] — specifically, this reduces the complexity of the problem from PSPACE-complete to PP-complete, where PP is a complexity class suspected to be smaller than PSPACE.

Assumptions of non-criticality allow the application of some very useful tools for the study of Gibbs states and ground states, namely quantum belief propagation [121] and the spectral flow [122]. In Section 4 we will show how these techniques can be used to obtain non-trivial many-to-one reductions of local Hamiltonian problems using classical postprocessing.

Though the assumption of non-criticality is physically-motivated and appears to be algorithmically useful, it is generally impossible to verify for a given family of Hamiltonians. This was established by Ref. [123], which showed that no algorithm can decide whether a family of translationally-invariant Hamiltonians in 2D remains spectrally gapped as $n \to \infty$.

Average-case hardness

One perspective on Hamiltonian complexity theory which, to this Author's knowledge, is very poorly understood (i.e. perhaps extremely difficult to say anything meaningful about) is the notion of average-case hardness. Since QMA-hard instances of the local Hamiltonian problem are rather contrived constructions, it is natural to ask whether there might exist algorithms (quantum or classical) which are successful when applied

to "most" instances, over some probability measure on the space of Hamiltonians. We can compare this to the case of classical 3-SAT, where it is known that the ratio of clauses m to the number of variables n is a determining factor in average-case hardness: for m/n below a constant, it is known that nearly all random instances are satisfiable, whilst for m/n above this constant nearly all random instances are unsatisfiable. Average-case hardness appears to appear in critical instances localised around the transition point $m/n \approx 4.2$ [124, 125]. Based on the above, average-case hardness of local Hamiltonian problems might require identifying regimes where criticality is unlikely.

Structural restrictions

Here we mention a few ways that the local Hamiltonian problem can be simplified by placing restrictions on the local terms h_A appearing in Eq. (2.5).

Firstly, one may require that each h_A is diagonal in the computational basis. This leads to the regime of classical Hamiltonians, for which finding ground state energies is known to be NP-complete for $k \geq 2$ [126]. Furthermore, requiring that the off-diagonal elements of h_A are non-positive leads to the family of so-called stoquastic Hamiltonians, which appear to avoid QMA-completeness [92] and are instead complete for a class known as StoqMA. These observations were generalised by Ref. [86], which classified the complexity of all local Hamiltonian problems where the h_A are drawn from finite set of interactions, and weighted by real numbers.

An alternative approach (also extending the classical case) is to assume that the h_A pairwise commute — this is known as the commuting local Hamiltonian problem, which was shown in Ref. [127] to be in NP for $k = 2$. Subsequent works have established NP containment also for $k = 3$ [128], and for 2D systems [129]. Whether the commuting local Hamiltonian problem is generally contained in NP remains an open question.

An additional variant on the local Hamiltonian problem, introduced by Ref. [130] is known as quantum k-SAT. Here, the terms h_A are each k-local projectors, and the problem is to decide whether or not the ground state is zero or at least $1/\operatorname{poly}(n)$, promised that one of these is the case. Ref. [130] showed that this problem is classically easy for $k = 2$, whilst Ref. [131] established that for $k \geq 3$ the problem is QMA_1 complete (for QMA_1 a complexity class which lies between NP and QMA).

Bounded interaction strengths

In the case of the 2-local Hamiltonian problem, all QMA-hard instances known to this Author are constructed via perturbative reductions, and therefore contain interaction strengths which range over many orders of magnitude. This is a manifestly non-physical feature which also extends to QMA-hard instances of the electronic structure problem, since these are obtained by perturbative reductions from 2-local Hamiltonians [85, 97]. We might therefore wonder whether QMA-hardness actually persists for 2-local Hamiltonians with bounded interaction strengths. One way to enforce this notion is to restrict to Hamiltonians built from a fixed (that is, not changing as n increases) set of 2-local interactions $\mathcal{S}$. We define this below.

Definition 2.4.1 ($\mathcal{S}$-bounded k-local Hamiltonian problem)**.** Given a finite set $\mathcal{S}$ consisting of k-local qubit interactions $\mathcal{S} \subseteq \mathrm{Herm}((\mathbb{C}^2)^{\otimes k})$, the $\mathcal{S}$-bounded Hamiltonian problem consists of instances of the k-local Hamiltonian for H of the form

$$H = \sum_{\substack{A \subseteq [n] \\ |A|=k}} h_A \ , \quad h_A \in \mathcal{S} \cup \{0\} \quad \text{for all } A \ . \tag{2.14}$$

Note that this is a different setting to the $\mathcal{S}$-Hamiltonian problem of Ref. [86], in which the authors allow the interactions to be scaled by arbitrary real numbers.

This problem is QMA-complete for $k \geq 3$: using a novel clock construction in the circuit-to-Hamiltonian mapping, the authors of Ref. [132] give a QMA-hard construction for 3-local Hamiltonians which does not use perturbative gadgets. Since we may assume that the QMA verification circuit is built from a finite gate set, this naturally consists of a finite set of interactions. Using the geometric layout of Ref. [72], this hardness can be seen to persist for instances containing bounded-range interactions in the 2D plane.

On the other hand, proving QMA-hardness for $k = 2$ and for any finite set $\mathcal{S}$ appears much trickier. For starters, circuit-to-Hamiltonian constructions face a fundamental barrier: generally they consist of terms which couple two qubits (on which a gate acts) with at least one clock qubit. It is therefore not clear how one could construct such an object restricted to 2-local terms or whether it should even be possible. Moreover, in Section 3 we will establish quite a general no-go result prohibiting locality reduction without strong interactions.

A couple of special cases of the $\mathcal{S}$-bounded 2-local Hamiltonian problem (in fact, for $|\,\mathcal{S}\,| = 1$) have known hardness:

- For $\mathcal{S} = \{Z \otimes Z\}$, we recover the classical Ising model, which is known to be NP-complete, even with restricted geometry [126].
- For $\mathcal{S} = \{X \otimes X + Y \otimes Y + Z \otimes Z\}$ this problem is known as quantum MAX-CUT and was recently shown to be NP-hard [133], though QMA-hardness remains an open question (see Ref. [60], Open Question 3.2).

Clearly the hardness of this problem will depend sensitively on the precise set $\mathcal{S}$ in question, but it is still not clear whether any finite set will give something QMA-complete for $k = 2$. Both k-SAT and QUANTUM-k-SAT exhibit large jumps in complexity between $k = 2$ and $k = 3$, although MAX-k-SAT, possibly the clearest classical analogue to the local Hamiltonian problem, does not [134]. As we note below, a positive resolution to the quantum PCP conjecture would give a finite set $\mathcal{S}$ for which QMA-hardness would hold, but this result would not extend to geometrically local Hamiltonians. The construction of Ref. [93] establishes a family of 2-local bosonic Hamiltonians which do maintain QMA-hardness even in the absence of strong interactions, however the qualitatively different bosonic statistics mean that these results do not immediately translate to the case of qubits.

The quantum PCP conjecture

It would be completely outrageous to conclude this section without mentioning perhaps the most famous open problem in quantum complexity theory: the quantum PCP conjecture [106, 96]. This is a quantum extension of the (classical) PCP theorem, a landmark result of computer science [64]. Roughly speaking, the quantum PCP conjecture states that the k-local Hamiltonian problem remains QMA-hard even up to constant relative precision, i.e. when the promise gap is bounded below as

$$\beta - \alpha \geq \epsilon n \ , \tag{2.15}$$

for some universal constant $\epsilon > 0$.

The quantum PCP conjecture would have consequences for the complexity of thermal states: since the state $\rho_\beta \sim e^{-\beta H}$ approximates the ground state up to error $\sim n/\beta$, a positive resolution to the conjecture would imply that ρ_β is QMA-hard to prepare even for constant β, extending the hardness regime of the problem given by Definition 2.2.5.

If the quantum PCP conjecture holds, then it holds for $k = 2$, due to an observation of Ref. [135]: to simulate ground state energies up to constant relative error, it is sufficient to use gadgets with constant interaction strengths. In fact, from here we could deduce QMA-hardness of the $\mathcal{S}$-bounded 2-local Hamiltonian problem (Definition 2.4.1) for some finite interaction set, by truncating all of the interactions in the subsequent Hamiltonian to some constant precision. It is important to note, however, that families of local Hamiltonians satisfying the conditions of the quantum PCP conjecture cannot be geometrically local. This follows from a straightforward argument: for any local Hamiltonian embedded in a D-dimensional lattice the ground state energy could be approximated to arbitrary relative precision by dividing the lattice into "patches" of volume $\sim \log n$ and solved locally; the error incurred by ignoring the boundaries of these patches is sublinear in n and thus does not close the promise gap. This limits the extent the extent to which the quantum PCP conjecture can be interpreted physically.

There are good reasons to believe that the quantum PCP conjecture is true, such as the recent proof of the weaker NLTS theorem [136, 137], and not least the classical PCP theorem [64]. On the other hand, structural restrictions on possible PCP Hamiltonians were given in Ref. [138] which indicate that a naive "quantisation" of the classical PCP theorem is probably not viable, and the likelihood of a positive resolution to the conjecture remains up for debate.

2.5 Consequences for quantum simulators

In this section, we have defined some fundamental computational problems arising in the simulation of quantum systems. One motivation for formalising these problems was to present convincing and rigorous arguments that they offer opportunities for practical applications of quantum computers, but the true picture turns out to be rather obscure. On one hand, worst-case hardness results suggest that quantum computers may be no more use than classical machines for general ground-state energy problems, which turn

Assumption	Known hardness results
General k-local Hamiltonian	$k \geq 2$, $D \geq 2$: QMA-complete [57, 72] $k \geq 2$, $D = 1$: QMA-complete [88, 89]
Good guiding state	$k \geq 2$, $D \geq 2$: BQP-complete [115]
Constant spectral gap	$D = 1$: in P [118]
Classical systems	$k \geq 2$, $D \geq 2$: NP-complete [126] $D = 1$: in P [118]
Commuting terms	$k = 2, 3$: NP-complete [127, 128] $D = 2$: NP-complete [129]
Quantum k-SAT	$k = 2$: in P [130] $k \geq 3$: QMA_1-complete [131]
Constant relative precision	$k \geq 2$: QMA-complete assuming qPCP [135]
Bounded interactions	$k \geq 2$: QMA-complete assuming qPCP

Figure 2.4: A summary of some of the hardness results mentioned in Section 2.4 for restricted k-local Hamiltonian problems on D-dimensional systems of qubits. "qPCP" stands for the quantum PCP conjecture.

out to be QMA-hard. On the other hand, it is unclear the extent to which worst-case complexity theory is relevant for real-life problems; typically hard instances are unphysical and pathological constructions, whilst real systems exhibit structure often admitting even good classical heuristic approaches. By studying restricted variants of the local Hamiltonian problem we can obtain insights into the sources of worst-case hardness and how it might be avoided, but it remains a major challenge in both theory and experiment to identify a concrete regime of practically useful quantum simulation tasks where quantum computers can operate beyond the reach of their classical counterparts.

To this end, the study of analogue quantum simulation is useful for two reasons. Naturally, from the practical perspective, analogue quantum simulators provide an accessible regime for near-term technologies to produce meaningful qualitative predictions and benchmarks for quantum systems. Secondly, from the theoretical perspective, understanding when one family of Hamiltonians can simulate another allows reductions between different simulation problems; determining when physically reasonable simulations are (and are not) possible gives insights into the types of problems we can expect to solve.

The remainder of this thesis will focus on analogue quantum simulation, and in particular attempts to avoid unrealistically strong interactions in perturbative simulations. In Section 3 (which is a reprint of Ref. [1]), the experimental and theoretical foundations of analogue quantum simulators will be discussed in more depth, followed by a general mathematical characterisation of modular gadgets for quantum simulation, and ultimately a no-go theorem for simulations without strong interactions. We will then show an example of how to circumvent the no-go theorem by considering a construction beyond pure Hamiltonian simulation, exploiting dissipative dynamics and the quantum Zeno effect. In Section 4 (a reprint of Ref. [2]), we will consider a quite different approach,

using classical postprocessing to effectively reduce the necessary interaction strengths for perturbative simulations.

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

The following section is a lightly edited version of the following article:

> [1]: Dylan Harley, Ishaun Datta, Frederik Ravn Klausen, Andreas Bluhm, Daniel Stilck França, Albert H. Werner, and Matthias Christandl. Going beyond gadgets: The importance of scalability for analogue quantum simulators. *Nature Communications*, 15(1):6527, 2024.

This article was written as part of an integrated MSc/PhD programme, and accordingly also appeared (in an edited form) as part of this Author's MSc thesis.

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.

3.1 Introduction

The simulation of quantum systems has long been identified as a potential application for quantum technologies [16], for which long-term benefits may range from condensed matter physics to quantum chemistry and the life sciences [139, 12]. 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 [43]. This approach, known as digital quantum simulation [140], has seen extensive theoretical development [44, 46] and remains a promising route towards attaining quantum advantage [30]. However, useful and scalable simulations remain out of reach for near-term technology [32] 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 [51]. 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 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.[56], 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 [141, 142, 143, 91, 144] 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. 3.1). Similar scalings are observed to arise through the use of Hamiltonian gadgets [72, 87], 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

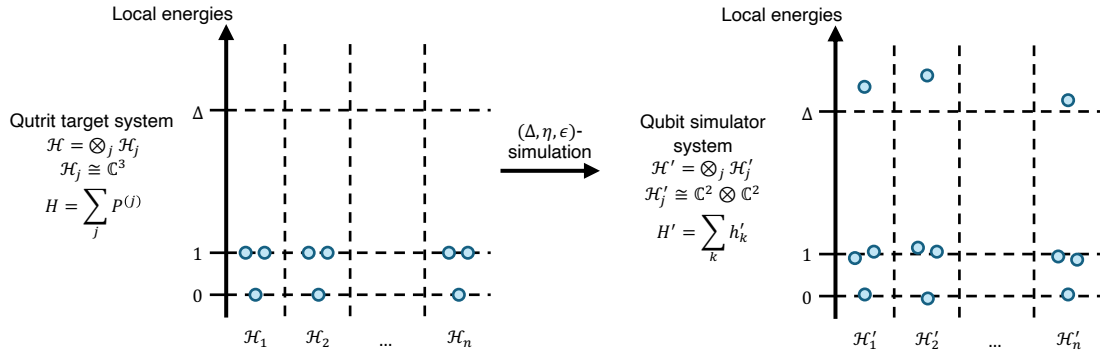


Figure 3.1: Qutrit-to-qubit encoding energies. A sketch of the on-site energies for a system of non-interacting qutrits under a low-energy encoding, before and after simulation in the sense introduced by Cubitt et al.[56]. In this example, each qutrit is mapped to a system of two qubits. The original 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 Δ 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.

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 [56] 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.[145]). Furthermore, as the experimental distinction between analogue and digital devices becomes increasingly blurred [32, 146], 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 [147, 148, 149].

3.2 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 Ω_{state}, and a set of observables Ω_{obs}; the goal of the simulator is then to approximate the evolution of the observables in Ω_{obs} under H, starting from initial states in Ω_{state}. Restricting the set of states Ω_{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 Ω_{state} to a specific invariant subspace and simulate only the reduced Hamiltonian. Likewise, Ω_{obs}

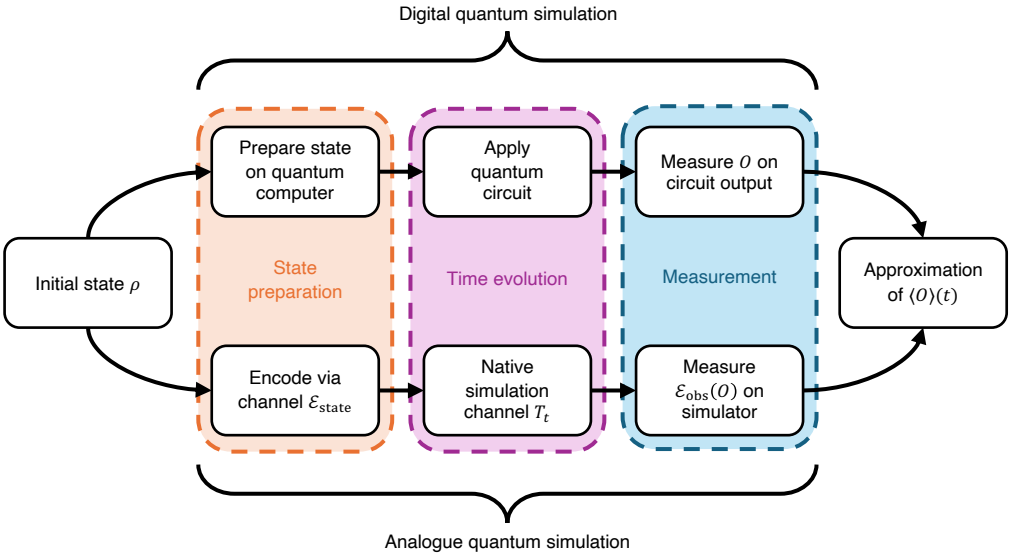


Figure 3.2: Analogue and digital simulation. A schematic description of our framework for analogue quantum simulation, in contrast with the digital approach. Both approaches aim to compute the observable expectation value $\langle O \rangle(t) = \mathrm{tr}[Oe^{-itH}\rho e^{itH}]$, given an initial state ρ. The analogue simulator, which uses state encodings $\mathcal{E}_{\text{state}}$ and $\mathcal{E}_{\text{obs}}$ respectively, has accuracy ϵ if $|\,\mathrm{tr}[Oe^{-itH}\rho e^{itH}] - \mathrm{tr}[\mathcal{E}_{\text{obs}}(O)(T_t \circ \mathcal{E}_{\text{state}})(\rho)]| \leq \epsilon$.

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 [54], significantly reducing the hardware overhead necessary to simulate for small times. Such techniques have been studied in the context of many-body state exploration [150, 151], and more recently in the realm of analogue simulation [55].

An analogue quantum simulator can be mathematically described in terms of three components, which we illustrate in Fig. 3.2. 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 time evolution is specified by a family of quantum channels $\{T_t\}_{t\in[0,t_{\max}]}$. 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.[51]), 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 $t \in [0, t_{\max}]$, 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 ϵ if the expectation value of this measurement is within ϵ 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 [152, 108].

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 [82, 72, 135, 87]. These provide a recipe to simulate complicated many-body interactions from a more restrictive family, for example to simulate a 3-body 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.[87] 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 Δ 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 [153, 154, 155] 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. 3.3(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. 3.3(a):

- Accuracy: The spectrum of H should be approximated by that of H', in some

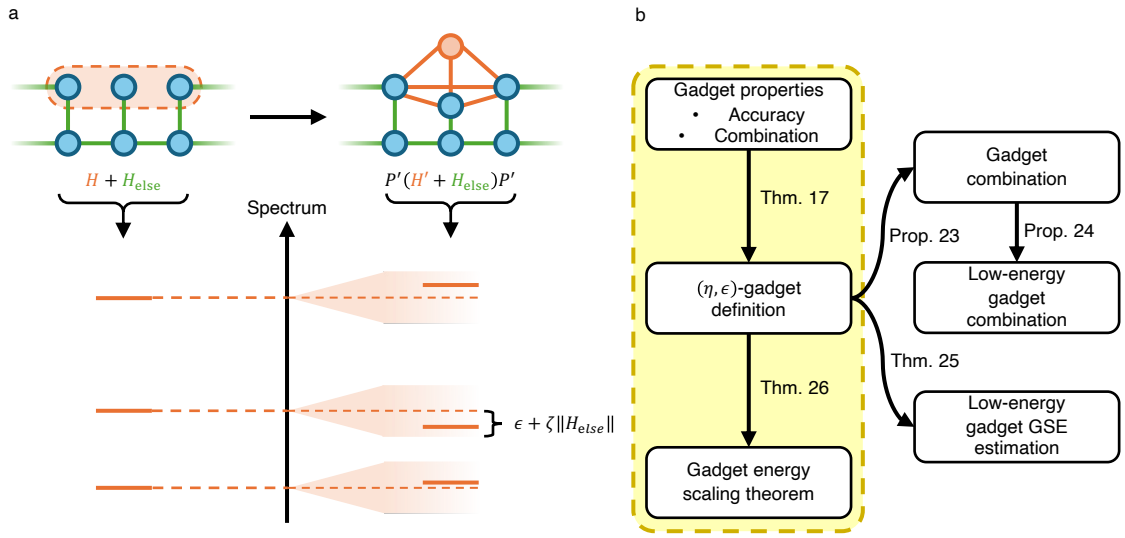


Figure 3.3: Hamiltonian gadget characterisation. (a) The interaction hypergraph of a Hamiltonian containing a 3-local interaction which is replaced by a 2-local gadget. The gadget property requires that the spectrum is unchanged up to $\epsilon + \zeta\|H_{\text{else}}\|$, for $\epsilon, \zeta > 0$, 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.[135] for ground state energy (GSE) estimation.

subspace defined by a projector $P' \in \mathrm{Proj}(\mathcal{H} \otimes \mathcal{A})$, up to error $\epsilon \geq 0$.

- Combination: The above property should hold even when any additional Hamiltonian H_{else} is added to both H and H', at the expense of an additional spectral error $\zeta \|H_{\mathrm{else}}\|$, where $\|\cdot\|$ denotes the operator norm.

For small error parameters ϵ and ζ, these properties are — non-trivially — sufficient to force H' to satisfy the following definition, which resembles previously studied notions of simulation [56, 87]. We say that H' is an (η, ϵ)-gadget (where η is a new parameter related to ζ, also measuring the ability of the gadget to combine with other terms) for H if there exists a projector $P \in \mathrm{Proj}(\mathcal{A}) \setminus \{0\}$ and a unitary operator $U \in \mathrm{U}(\mathcal{H} \otimes A)$ satisfying two conditions. Firstly, U must be η-close to the identity: $\|U - \mathcal{I}\| \leq \eta$. Then, defining the projector P' by $P' = U(\mathcal{I} \otimes P)U^\dagger$ (so that η 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: ϵ can be thought of as the absolute error of the gadget, whilst η bounds the error incurred when the gadget is combined with other interactions in a Hamiltonian. In particular, when $\|H_{\mathrm{else}}\| \sim n$ grows with the size of the system, η 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.[135].

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, η^{-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.

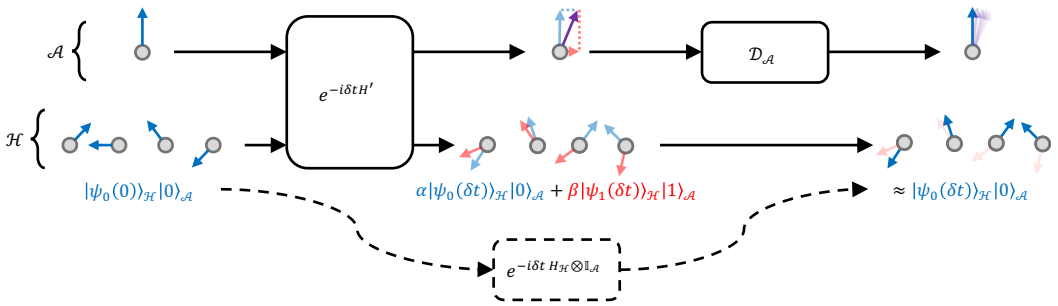


Figure 3.4: Dissipative gadget evolution. Circuit representation of the non-unitary gadget procedure for a single timestep. The initial state $|\psi_0(0)\rangle_{\mathcal{H}}|0\rangle_{\mathcal{A}}$ is evolved under a Hamiltonian H' for time δt, resulting in a superposition of states with the ancillary 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, $H_{\mathcal{H}}$.

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 Δ scaling with the size of the system. Instead, a dissipative channel is applied to the qubit m at regular intervals separated by time δt. Provided that δt 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. 3.4, and a rigorous description can be found in the Methods.

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[156, 157] and in practice[158] 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

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 $\|000\ldots\rangle$.	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 3.5: A summary of the DiVincenzo[159] and Cirac-Zoller[51] 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.

dynamics, we anticipate that similar hybrid techniques will constitute a powerful tool for attaining useful quantum advantage with quantum simulators.

3.3 Analogue quantum simulation

3.3.1 Criteria for quantum computation and simulation

In a review of the prospective possibilities of quantum computing [159] 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. 3.5.

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.

A similar set of criteria for analogue quantum simulators is discussed by Cirac et al.[51], also summarised in Fig. 3.5. 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 $\mathrm{U}((\mathbb{C}^2)^{\otimes k})$ of unitary

transformations relatively cheaply due to the Solovay-Kitaev theorem [160], the task of an analogue quantum simulator is to implement k-qudit interactions from the unbounded set of possible Hamiltonians $\mathrm{Herm}((\mathbb{C}^d)^{\otimes k})$. 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 3.3.1 (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 $\Theta(1)$.

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 3.3.1. 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.[161].

As well as being experimentally and qualitatively desirable, encoding interactions independently has quantitative benefits; as noted by Cubitt et al.[56], 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.[145] show that restriction to

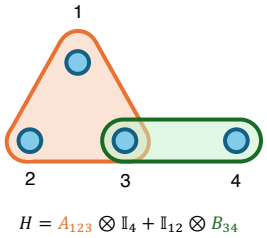


Figure 3.6: Example Hamiltonian interaction hypergraph. A Hamiltonian H on 4 qubits, and its associated interaction hypergraph. The Hamiltonian consists of a 3-local (orange) and a 2-local (red) term, so we say that H is 3-local.

such Hamiltonians will necessarily sacrifice some sense of universality of the simulator. Earlier results in Hamiltonian complexity theory[135], however, show that in many cases it is still possible to simulate ground state energies up to an extensive error.

3.3.2 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. 3.6.

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 3.3.2 (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 [39], it was established that the 5-local Hamiltonian problem is QMA-complete, and subsequent works

optimising the construction [82] and using gadget techniques [57] 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 [89, 86]); indeed hardness results have been shown to hold even under the significant restriction to 1-dimensional translationally invariant systems [90]. QMA-completeness is closely related to a notion of universality for simulators; an equivalence was proved by Kohler et al.[91].

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.[135], 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.[56], in which the authors study methods of encoding Hamiltonians via a map $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \to \text{Herm}(\mathcal{H}')$, 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 3.3.3 (Local Hamiltonian encoding [56])**.** A local Hamiltonian encoding is a map $\mathcal{E}_{\text{obs}} : \text{Lin}(\otimes_{i=1}^{n} \mathcal{H}_i) \to \text{Lin}(\otimes_{i=1}^{n} \mathcal{H}'_i)$ of the form

$$\mathcal{E}_{\text{obs}}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger \ , \tag{3.1}$$

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 $(P_i \otimes \mathcal{I})P = P$ and $(Q_i \otimes \mathcal{I})Q = Q$. Generally, we consider the case of $\text{rank}(P) > 0$ (referred to as standard[56]), for which one can define a corresponding state encoding

$$\mathcal{E}_{\text{state}}(\rho) = V(\rho \otimes \tau)V^\dagger \ , \tag{3.2}$$

where τ 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 3.3.4 ((Δ, η, ϵ)-simulation [56])**.** A Hamiltonian $H' \in \mathrm{Herm}(\mathcal{H}') = \mathrm{Herm}(\otimes_{i=1}^{n} \mathcal{H}'_i)$ is said to (Δ, η, ϵ)-simulate a Hamiltonian $H \in \mathrm{Herm}(\mathcal{H}) = \mathrm{Herm}(\otimes_{i=1}^{n} \mathcal{H}_i)$ if there exists a local encoding (Definition 3.3.3) $\mathcal{E}_{\mathrm{obs}}(M) = V(M \otimes P + \bar{M} \otimes Q)V^\dagger$ such that

(i) There exists an encoding $\tilde{\mathcal{E}}_{\mathrm{obs}}(M) = \tilde{V}(M \otimes P + \bar{M} \otimes Q)\tilde{V}^\dagger$ (where $\tilde{V} \in \mathrm{Isom}(\mathcal{H} \otimes \mathcal{A}, \mathcal{H}')$ need not have a tensor product structure as in Definition 3.3.3) such that $\|V - \tilde{V}\| \leq \eta$ and $\tilde{\mathcal{E}}_{\mathrm{obs}}(\mathcal{I}) = P_{\leq \Delta(H')}$ is the projection onto the low-energy ($\leq \Delta$) subspace of H', and

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

This approach (later generalised by Apel et al.[162] and refined with resource constraints by Zhou et al.[144]) 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 3.3.1.

3.3.3 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 $t_{\max}$. 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 Ω_{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 Ω_{obs}. We denote by $\mathcal{H}'$ the Hilbert space corresponding to the simulator system, and for $t \in [0, t_{\max}]$ we write $T_t : D(\mathcal{H}') \to D(\mathcal{H}')$ 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 [150, 151, 55].

The minimal requirement for a simulator is that it should approximate the expectation values of the elements of Ω_{obs}. That is, $\mathrm{tr}[Oe^{-iHt}\rho e^{iHt}]$ should be close to $\mathrm{tr}[O'T_t(\rho')]$ for all $\rho \in \Omega_{\mathrm{state}}$ and $O \in \Omega_{\mathrm{obs}}$, where ρ' 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 ρ and O, with $\mathcal{H}'$ a space large enough to contain all of them and by encoding ρ into several copies. However, this would violate the size-independence requirement of Definition 3.3.1 if Ω_{obs} and Ω_{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 $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \to D(\mathcal{H}')$. Correspondingly, to accommodate for quantum outputs, we require the observable encoding $O \mapsto O'$ to be a unital and completely positive map $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \to \text{Herm}(\mathcal{H}')$. 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 ρ' 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 [163]), 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 [164]), and summing the average results.

The above discussion leads us to the following definition, which is illustrated by Fig. 3.2.

Definition 3.3.5 (Analogue quantum simulation)**.** Given a set of states Ω_{state} on a Hilbert space $\mathcal{H}$, a normalised set of observables Ω_{obs} (i.e. $\|O\| = 1$ for all $O \in \Omega_{\text{obs}}$, where $\|\cdot\|$ denotes the operator norm), a time $t_{\max} > 0$, a Hamiltonian $H \in \text{Herm}(\mathcal{H})$, and $\epsilon > 0$, we say that a family of quantum channels $T_t : D(\mathcal{H}') \to D(\mathcal{H}')$, for $t \in [0, t_{\max}]$ simulates H with respect to Ω_{state} and Ω_{obs} with accuracy ϵ if there exists

1. A state encoding quantum channel $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \to D(\mathcal{H}')$ which maps states to the simulator Hilbert space $\mathcal{H}'$,

2. An observable encoding, given by a unital and completely positive map $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \to \text{Herm}(\mathcal{H}')$,

such that

$$\left| \text{tr}[\mathcal{E}_{\text{obs}}(O)(T_t \circ \mathcal{E}_{\text{state}})(\rho)] - \text{tr}[O(e^{-itH}\rho e^{itH})]] \right| \leq \epsilon \ , \tag{3.3}$$

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

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 [152, 108]. It should be noted also that Definition 3.3.5 could equivalently have been phrased in terms of a set of POVMs rather than observables Ω_{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 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, (3.3) holds for any convex combination of the states and observables in Ω_{state} and Ω_{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 3.3.5 are taken simply as time evolution under some simulator Hamiltonian $H' \in \mathrm{Herm}(\mathcal{H}')$, 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 [165]. 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. 3.5. Despite the generality afforded by Definition 3.3.5, we emphasise that experimentally practical simulations should be size-independent as in Definition 3.3.1. 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[74] and previous notions of dynamical Hamiltonian simulation[166] 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 3.3.1.

3.3.4 Local encodings

Although Definition 3.3.5 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 Section 3.6.1. A similar discussion is presented for the stronger case of local Hamiltonian encodings of Cubitt et al.[56], and a discussion of the stability of local observable measurements to local noise is given by Trivedi et al.[55].

Definition 3.3.6 (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 $\mathcal{E}_{\mathrm{state}} : D(\mathcal{H}) \to D(\mathcal{H}')$ is local if it has a Stinespring representation of the form

$$\mathcal{E}_{\mathrm{state}}(\rho) = \mathrm{tr}_E[U(\rho \otimes |0\rangle\langle 0|_F)U^\dagger] \ , \tag{3.4}$$

where $F = \otimes_k F_k$ and $E = \otimes_l E_l$ are ancillary systems and $U \in \mathrm{U}(\mathcal{H} \otimes F, \mathcal{H}' \otimes E)$ 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[167].

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 $\mathcal{N}' : D(\mathcal{H}') \to D(\mathcal{H}')$ on the simulator, there exists a corresponding local error channel $\mathcal{N} : D(\mathcal{H}) \to D(\mathcal{H})$ on the target system satisfying

$$\mathcal{N}' \circ \mathcal{E}_{\text{state}}(\rho) \stackrel{?}{=} \mathcal{E}_{\text{state}} \circ \mathcal{N}(\rho) \ . \tag{3.5}$$

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 3.3.7 (Local error back-propogation)**.** *Let $\mathcal{E}_{\text{state}} : D(\mathcal{H}) \to D(\mathcal{H}')$ be a local state encoding as in Definition 3.3.6, and let $\mathcal{N}' : D(\mathcal{H}') \to D(\mathcal{H}')$ be a channel whose Kraus operators $\{X'_k\}$ each act on $O(1)$ sites in $\mathcal{H}'$. Then there exists a channel $\mathcal{N} : D(\mathcal{H} \otimes F) \to D(\mathcal{H} \otimes F)$ whose Kraus operators $\{X_k\}$ each act on $O(1)$ sites in $\mathcal{H} \otimes F$, and such that for all $\rho \in \mathcal{H}$,*

$$\mathcal{N}' \circ \mathcal{E}_{\text{state}}(\rho) = \text{tr}_E[U\mathcal{N}(\rho \otimes |0\rangle\langle 0|_F)U^\dagger] \ . \tag{3.6}$$

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-propogation 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 3.3.8 (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 $\mathcal{E}_{\text{obs}} : \text{Herm}(\mathcal{H}) \to \text{Herm}(\mathcal{H}')$ 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 $\mathcal{E}^*_{\text{obs}} : D(\mathcal{H}') \to D(\mathcal{H})$ 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 3.3.9 (Encoded measurements)**.** *Let $\mathcal{E}_{\text{obs}}$ be a local observable encoding as in Definition 3.3.8, and let O be a local operator on $\mathcal{H}$. Then $\mathcal{E}_{\text{obs}}(O)$ can be measured using a local POVM on $\mathcal{H}'$.*

3.3.5 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 3.3.5. 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 3.3.4. 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.[56] (Definition 3.3.4) 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 3.3.10 provides a formal statement of this fact (proved in Section 3.6.2) 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 jth qutrit.

Proposition 3.3.10. *Let $\mathcal{H} = (\mathbb{C}^3)^{\otimes n}$ be the space of n qutrits acted on by the Hamiltonian*

$$H_n = \sum_{j=1}^{n} (P_0^{(j)} + P_\uparrow^{(j)}) \ . \tag{3.7}$$

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 (Δ, η, ϵ)-simulation for H_n in the sense of Definition 3.3.4, for $\eta \in [0, 1)$ and $\epsilon \geq 0$, then

$$\max_j \|h'_j\| = \Omega(n^{1-\alpha}(1-\eta^2)) \ . \tag{3.8}$$

From (3.8) 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 3.3.1 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. 3.1. In contrast, the simulation task is trivial in our framework given in Definition 3.3.5 because the low-energy encoding requirement is relaxed.

Letting $H_n = \sum_{j=1}^{n}(P_0^{(j)} + P_\uparrow^{(j)})$ as in Proposition 3.3.10, 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

$$V : \mathbb{C}^3 \to \mathbb{C}^2 \otimes \mathbb{C}^2 \ , \tag{3.9}$$

encoding each qutrit into two qubits. To realise a simulator in the sense of Definition 3.3.5, we let

$$\mathcal{E}_{\text{state}} : \rho \mapsto V^{\otimes n}\rho(V^{\otimes n})^\dagger \ , \quad \mathcal{E}_{\text{obs}} : O \mapsto V^{\otimes n}O(V^{\otimes n})^\dagger \ , \tag{3.10}$$

and

$$T_t = e^{-it\mathcal{E}_{\text{obs}}(H_n)}(\cdot)e^{it\mathcal{E}_{\text{obs}}(H_n)} \ , \tag{3.11}$$

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

Although Proposition 3.3.10 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 3.3.10 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 3.3.11. *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 $d(x, y) : \mathbb{Z}^D \times \mathbb{Z}^D \to \mathbb{Z}$ 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

$$|B_R(y)| := |\{X_i : \exists x \in X_i \text{ with } d(x, y) \leq R\}| = O\big(n^{1-\min\{\alpha,1/D\}}\big) \ . \tag{3.12}$$

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[168, 169].

Noisy analogue simulators Suppose we have quantum channels T_t, for $t \in [0, t_{\max}]$ which simulate some $H \in \mathrm{Herm}(\mathcal{H})$ with respect to Ω_{state} and Ω_{obs} up to accuracy ϵ as in Definition 3.3.5, corresponding to encoding maps $\mathcal{E}_{\mathrm{state}}$ and $\mathcal{E}_{\mathrm{obs}}$.

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 $\tilde{T}_t$, $\tilde{\mathcal{E}}_{\mathrm{state}}$, and $\tilde{\mathcal{E}}_{\mathrm{obs}}$. For any $O \in \Omega_{\mathrm{obs}}$, $\rho \in \Omega_{\mathrm{state}}$, we may bound the additional error in observable expectation values incurred by the noisy maps by

$$\begin{aligned}
&|\operatorname{tr}[\mathcal{E}_{\mathrm{obs}}(O)(T_t \circ \mathcal{E}_{\mathrm{state}})(\rho)] - \operatorname{tr}[\tilde{\mathcal{E}}_{\mathrm{obs}}(O)(\tilde{T}_t \circ \tilde{\mathcal{E}}_{\mathrm{state}}(\rho)]| \\
&\quad = |\operatorname{tr}[O\big(\mathcal{E}^*_{\mathrm{obs}} \circ T_t \circ \mathcal{E}_{\mathrm{state}} - \tilde{\mathcal{E}}^*_{\mathrm{obs}} \circ \tilde{T}_t \circ \tilde{\mathcal{E}}_{\mathrm{state}}\big)(\rho)] \\
&\quad \leq \|\mathcal{E}^*_{\mathrm{obs}} \circ T_t \circ \mathcal{E}_{\mathrm{state}} - \tilde{\mathcal{E}}^*_{\mathrm{obs}} \circ \tilde{T}_t \circ \tilde{\mathcal{E}}_{\mathrm{state}}\|_{1\to 1} \\
&\quad \leq \|\mathcal{E}^*_{\mathrm{obs}} - \tilde{\mathcal{E}}^*_{\mathrm{obs}}\|_{1\to 1} + \|T_t - \tilde{T}_t\|_{1\to 1} + \|\mathcal{E}_{\mathrm{state}} - \tilde{\mathcal{E}}_{\mathrm{state}}\|_{1\to 1} \ , \qquad (3.13)
\end{aligned}$$

where $\|\cdot\|_{1\to 1}$ denotes the one-to-one norm $\|\Lambda\|_{1\to 1} = \sup_\rho \|\Lambda(\rho)\|_1$ (defined as the induced trace norm[170] — 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 Ω_{state} and Ω_{obs}, up to error

$$\epsilon' \leq \epsilon + \sup_t \|T_t - \tilde{T}_t\|_{1\to 1} + \|\mathcal{E}_{\mathrm{state}} - \tilde{\mathcal{E}}_{\mathrm{state}}\|_{1\to 1} + \|\mathcal{E}^*_{\mathrm{obs}} - \tilde{\mathcal{E}}^*_{\mathrm{obs}}\|_{1\to 1} \ . \qquad (3.14)$$

Local Hamiltonian simulation in a subspace Suppose that H' is a (Δ, η, ϵ)-simulation of H as defined by Cubitt et al.[56] (Definition 3.3.4), corresponding to encodings $\mathcal{E}_{\mathrm{state}}$ and $\mathcal{E}_{\mathrm{obs}}$, 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 3.3.5.

We make use of the following lemmas. Lemma 3.3.12 ensures that measurement and time evolution are consistent with the encodings of Definition 3.3.4, and Lemma 3.3.13 bounds the error of (Δ, η, ϵ)-simulations under time evolution.

Lemma 3.3.12 (Cubitt et al., Proposition 4[56])**.** *If $\mathcal{E}_{\mathrm{state}}$ and $\mathcal{E}_{\mathrm{obs}}$ are encodings as in Definition 3.3.4 and* (3.2)*, then for all observables O and states ρ on the target system $\mathcal{H}$,*

$$\operatorname{tr}[\mathcal{E}_{\mathrm{obs}}(O)\mathcal{E}_{\mathrm{state}}(\rho)] = \operatorname{tr}[O\rho] \ . \qquad (3.15)$$

Moreover if the encoding is standard ($\operatorname{rank}(P) > 0$ *in Definition 3.3.4) then*

$$e^{-i\mathcal{E}_{\mathrm{obs}}(H)t}\mathcal{E}_{\mathrm{state}}(\rho)e^{i\mathcal{E}_{\mathrm{obs}}(H)t} = \mathcal{E}_{\mathrm{state}}\big(e^{-iHt}\rho e^{iHt}\big) \ . \qquad (3.16)$$

Lemma 3.3.13 (Cubitt et al., Proposition 28[56])**.** *Let H' be a (Δ, η, ϵ)-simulation of H in the sense of Definition 3.3.4 corresponding to encodings $\mathcal{E}_{\mathrm{obs}}$, $\mathcal{E}_{\mathrm{state}}$. If ρ' is a state in the simulator system $\mathcal{H}'$ satisfying $\mathcal{E}_{\mathrm{obs}}(\mathcal{I})\rho' = \rho'$, then for all t*

$$\|e^{-iH't}\rho' e^{iH't} - e^{-i\mathcal{E}_{\mathrm{obs}}(H)t}\rho' e^{i\mathcal{E}_{\mathrm{obs}}(H)t}\|_1 \leq 2\epsilon t + 4\eta \ . \qquad (3.17)$$

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

$$\begin{aligned}
&|\operatorname{tr}[\mathcal{E}_{\text{obs}}(O)e^{-iH't}\mathcal{E}_{\text{state}}(\rho)e^{iH't}] - \operatorname{tr}[Oe^{-iHt}\rho e^{iHt}]| \\
&\quad = |\operatorname{tr}[\mathcal{E}_{\text{obs}}(O)\Big(e^{-iH't}\mathcal{E}_{\text{state}}(\rho)e^{iH't} - e^{-i\mathcal{E}_{\text{obs}}(H)t}\mathcal{E}_{\text{state}}(\rho)e^{i\mathcal{E}_{\text{obs}}(H)t}\Big)]| \\
&\quad \leq \|O\|(2\epsilon t + 4\eta) \ . \qquad (3.18)
\end{aligned}$$

Hence the channels $T_t : \rho' \mapsto e^{-iH't}\rho' e^{iH't}$, for $t \in [0, t_{\max}]$ simulate H in the sense of Definition 3.3.5 with respect to any Ω_{state} and Ω_{obs}, up to error

$$\epsilon' \leq 2\epsilon t_{\max} + 4\eta. \qquad (3.19)$$

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.[56] constitutes an analogue quantum simulator in our framework given by Definition 3.3.5.

Short-time simulation with Lieb-Robinson bounds One advantage of only requiring the simulation of a particular set of observables Ω_{obs} in Definition 3.3.5, as opposed to reproducing the entire physical system, is that one can take advantage of the limited spread of correlations for short-time dynamics [54]. 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 [150, 151], and more recently in the setting of analogue simulators [55]. 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

$$H_n = \sum_{x=1}^{n} h_x \ , \qquad (3.20)$$

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 Ω_{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. 3.7. This situation is studied by Trivedi et al.[55], in particular for the thermodynamic limit $n \to \infty$.

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 ρ and O simply by restricting them to the smaller subsystem. Then a simulation of an observable $O \in \Omega_{\text{obs}}$ up to accuracy ϵ, satisfying

$$|\operatorname{tr}[Oe^{-iH_nt}\rho e^{iH_nt}] - \operatorname{tr}[\mathcal{E}_{\text{obs}}(O)e^{-iH_mt}\mathcal{E}_{\text{state}}(\rho)e^{iH_mt}]| \leq \epsilon \ , \qquad (3.21)$$

can be accomplished in the large-n regime for all $t \in [0, t_{\max}]$ if one takes $m = O\big(\log^d(1/\epsilon) + t_{\max}^d\big)$ (see Trivedi et al.[55], Lemma 1).

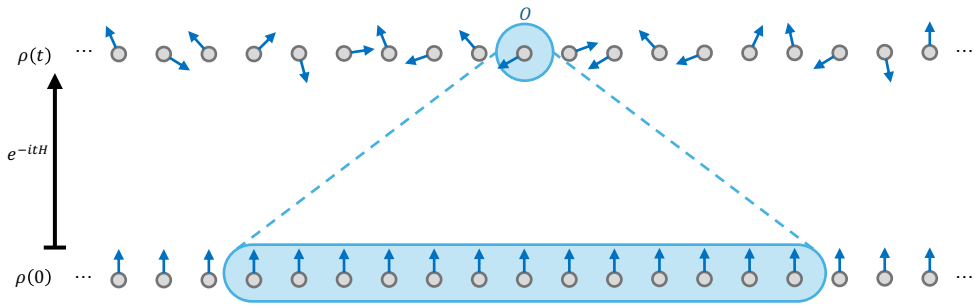


Figure 3.7: Simulation with Lieb-Robinson bounds. Simulation of a 1-dimensional spin system under a Hamiltonian H for time t. In theory, the system extends infinitely, but to estimate the value of a local observable O it is only necessary to simulate a subsystem corresponding to the Lieb-Robinson light cone.

3.4 Modular encodings and gadgets

3.4.1 General definitions

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 3.3.1, 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.[57] as a means of proving QMA-completeness of the 2-local Hamiltonian problem by reduction from the 3-local case [82], 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 [72] and in general to reduce Hamiltonians to more restrictive families of interactions [84, 85, 86]. Moreover, beyond Hamiltonian complexity-theoretic results, gadgets can be tailored to improve the performance of variational quantum algorithms [155].

In this work, we introduce a formalism which we argue encompasses any attempt at gadgetisation, in a sense which we make precise (Definition 3.4.1), in order to prove general properties of such constructions. Note that our approach, and the (η, ϵ) accuracy parameters, are closely related to those used in other definitions of simulation [56, 87]. 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 $H \in \mathrm{Herm}(\mathcal{H})$ 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 $H' \in \mathrm{Herm}(\mathcal{H} \otimes \mathcal{A})$, which acts on $O(1)$ sites in $\mathcal{H}$ and $\mathcal{A}$.

A simulator Hamiltonian in the sense of Definition 3.3.5 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 $P' \in \mathrm{Proj}(\mathcal{H} \otimes \mathcal{A})$ such that $P'(H' + H_{\mathrm{else}} \otimes \mathcal{I})P'$ approximates the spectrum of $H + H_{\mathrm{else}}$, for any $H_{\mathrm{else}} \in \mathrm{Herm}(\mathcal{H})$ (see Fig. 3.3(a)). We formalise this with the following definition.

Definition 3.4.1 ((ζ, ϵ)-gadget property)**.** Given a Hamiltonian $H \in \mathrm{Herm}(\mathcal{H})$ acting on a system $\mathcal{H} = \otimes_{i=1}^{n} \mathcal{H}_i$, and $H' \in \mathrm{Herm}(\mathcal{H} \otimes \mathcal{A})$ for $\mathcal{A}$ an ancillary system, we say that $(H', \mathcal{A})$ satisfies the (ζ, ϵ)-gadget property for H if there exists $P' \in \mathrm{Proj}(\mathcal{H} \otimes \mathcal{A})$, $\tilde{P} \in \mathrm{Proj}(\mathcal{A}) \setminus \{0\}$ such that, for any $H_{\mathrm{else}} \in \mathrm{Herm}(\mathcal{H})$, there exists a unitary $\tilde{U}_{H_{\mathrm{else}}} \in \mathrm{U}(\mathcal{H} \otimes \mathcal{A})$ with

$$\|P'(H' + H_{\mathrm{else}} \otimes \mathcal{I})P' - \tilde{U}_{H_{\mathrm{else}}}\big((H + H_{\mathrm{else}}) \otimes \tilde{P}\big)\tilde{U}_{H_{\mathrm{else}}}^{\dagger}\| \leq \epsilon + \zeta\|H_{\mathrm{else}}\| \ . \tag{3.22}$$

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

As noted by Cubitt et al.[56], 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 3.4.1 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_{else} in Definition 3.4.1 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 3.4.1 is a natural requirement, it is not convenient to work with due to the appearance of the general H_{else} acting on the entire of $\mathcal{H}$, upon which $\tilde{U}$ depends. The following alternative definition does not suffer from this problem.

Definition 3.4.2 ((η, ϵ)-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 (η, ϵ)-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

$$\|U - \mathcal{I}\| \leq \eta \ , \quad \|P'H'P' - U(H \otimes P)U^\dagger\| \leq \epsilon \ , \tag{3.23}$$

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

The advantage of Definition 3.4.2 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 3.4.1, 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 3.4.2, 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 Section 3.6.3.

Theorem 3.4.3 ((η, ϵ)-gadgets have the (ζ, ϵ)-gadget property)**.** *Suppose that $(H', \mathcal{A})$ is a (η, ϵ)-gadget for H. Then $(H', \mathcal{A})$ satisfies the (ζ, ϵ)-gadget property for H, where $\zeta = O(\eta)$.*

Theorem 3.4.4 (The (ζ, ϵ)-gadget property requires a (η, ϵ)-gadget)**.** *Suppose that $(H', \mathcal{A})$ satisfies the (ζ, ϵ)-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 (η, ϵ')-gadget for H, where $\eta = O(\epsilon) + O(\zeta^{\frac{1}{2}})$ and $\epsilon' = O(\epsilon) + O(\zeta)$.*

The roles of the η and ϵ parameters are to bound the error in the eigenvectors and eigenvalues respectively. Roughly speaking, η quantifies how well the gadget combines with other terms, and ϵ 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 η error.

Prior work in Hamiltonian complexity theory has focused on gadgetisation in the context of ground state estimation [82, 86, 87] or simulation in a low energy subspace [56]; 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 3.4.5 ((Δ, η, ϵ)-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 (Δ, η, ϵ)-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 $P_{\leq\Delta(H')} = U(\mathcal{I}\otimes P)U^\dagger$, and

$$\|U - \mathcal{I}\| \leq \eta \ , \quad \|P_{\leq\Delta(H')}H'P_{\leq\Delta(H')} - U(H\otimes P)U^\dagger\| \leq \epsilon \ . \tag{3.24}$$

In other words, the pair $(H', \mathcal{A})$ satisfy Definition 3.4.2, in the special case where we can use $P' = P_{\leq\Delta(H')}$.

Notice that Definition 3.4.5 imposes a significantly stronger requirement on H' than Definition 3.4.2; a priori there is no reason to expect that there will exist any choice of P and U such that $P_{\leq\Delta(H')} = U(\mathcal{I}\otimes P)U^\dagger$. Definitions 3.4.2 and 3.4.5 are sufficient to guarantee desirable combination properties, and are satisfied by widely-used constructions.

3.4.2 Examples of gadgets

Lemmas 4-7 of Bravyi et al.[87] can be naturally adapted to give several constructions for (Δ, η, ϵ) gadgets, which we use to demonstrate that Definition 3.4.2 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

$$V = \begin{pmatrix} V_{00} & V_{01} \\ V_{10} & V_{11} \end{pmatrix} \ , \tag{3.25}$$

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

Lemma 3.4.6 (First-order gadgets, adapted from Bravyi et al.[87])**.** *Suppose $H \in \mathrm{Herm}(\mathcal{H})$ and $V \in \mathrm{Herm}(\mathcal{H}')$ are such that*

$$\|H - V_{00}\| \leq \frac{\epsilon}{2} \ . \tag{3.26}$$

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

Lemma 3.4.7 (Second-order gadgets, adapted from Bravyi et al.[87])**.** *Let $H \in \mathrm{Herm}(\mathcal{H})$, and suppose $V^{(1)}, V^{(0)} \in \mathrm{Herm}(\mathcal{H}')$ are such that $\|V^{(1)}\|, \|V^{(0)}\| \leq \Lambda$, $V^{(0)}_{10} = V^{(0)}_{01} = V^{(1)}_{00} = 0$, and*

$$\|H - V^{(0)}_{00} + V^{(1)}_{01}V^{(1)}_{10}\| \leq \frac{\epsilon}{2} \ . \tag{3.27}$$

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

$$\Delta \geq O(\epsilon^{-2}\Lambda^6 + \eta^{-2}\Lambda^2) \ . \tag{3.28}$$

Lemma 3.4.8 (Third-order gadgets, adapted from Bravyi et al.[87])**.** *Let $H \in \mathrm{Herm}(\mathcal{H})$, and suppose $V^{(2)}, V^{(1)}, V^{(0)} \in \mathrm{Herm}(\mathcal{H}')$ are such that $\|V^{(2)}\|, \|V^{(1)}\|, \|V^{(0)}\| \leq \Lambda$,*

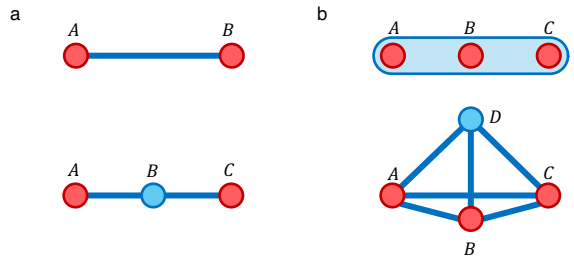


Figure 3.8: Existing gadgets. (a) Interaction hypergraphs of a 2-system interaction before (above) and after (below) the use of the subdivision gadget. (b) Interaction hypergraphs of a 3-system interaction before (above) and after (below) the use of the 3-to-2 gadget.

$$V_{10}^{(1)} = V_{01}^{(1)} = V_{10}^{(0)} = V_{01}^{(0)} = 0,\ V_{00}^{(2)} = 0,$$

$$\|H - V_{00}^{(0)} - V_{01}^{(2)} V_{11}^{(2)} V_{10}^{(2)}\| \leq \frac{\epsilon}{2}\ , \quad \textit{and} \quad V_{00}^{(1)} = V_{01}^{(2)} V_{10}^{(2)}\ . \tag{3.29}$$

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* $H_0 = \mathcal{I} \otimes |1\rangle\langle 1|$*, if*

$$\Delta \geq O(\epsilon^{-3}\Lambda^{12} + \eta^{-3}\Lambda^3)\ . \tag{3.30}$$

We illustrate the application of these lemmas to our definition with the following ubiquitous gadgets from Oliviera et al.[72]:

Given a target Hamiltonian $H = A \otimes B \in \mathrm{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B)$, 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

$$H' = \Delta H_0 + \Delta^{\frac{1}{2}} V^{(1)} + V^{(0)}\ , \tag{3.31}$$

where

$$H_0 = \mathcal{I} \otimes \mathcal{I} \otimes |1\rangle\langle 1|\ , \tag{3.32}$$

$$V^{(1)} = \frac{1}{\sqrt{2}}(-A \otimes \mathcal{I} + \mathcal{I} \otimes B) \otimes X\ , \tag{3.33}$$

$$V^{(0)} = \frac{1}{2}(A^2 \otimes \mathcal{I} + \mathcal{I} \otimes B^2) \otimes \mathcal{I}\ . \tag{3.34}$$

Then by Lemma 3.4.7 we see that, for sufficiently large Δ, $(H', \mathcal{H}_C)$ defines a $(O(\Delta), \eta, \epsilon)$-gadget for H (see Fig. 3.8(a)).

Given a target Hamiltonian $H = A \otimes B \otimes C \in \mathrm{Herm}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, 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

$$H' = \Delta H_0 + \Delta^{\frac{2}{3}} V^{(2)} + \Delta^{\frac{1}{3}} V^{(1)} + V^{(0)}\ , \tag{3.35}$$

where

$$H_0 = \mathcal{I}\otimes\mathcal{I}\otimes\mathcal{I}\otimes|1\rangle\langle 1| \ , \tag{3.36}$$

$$V^{(2)} = \frac{1}{\sqrt{2}}(-A\otimes\mathcal{I}+\mathcal{I}\otimes B)\otimes\mathcal{I}\otimes X - \mathcal{I}\otimes\mathcal{I}\otimes C\otimes|1\rangle\langle 1| \ , \tag{3.37}$$

$$V^{(1)} = \frac{1}{2}(-A\otimes\mathcal{I}+\mathcal{I}\otimes B)^2\otimes\mathcal{I}\otimes\mathcal{I} \ , \tag{3.38}$$

$$V^{(0)} = \frac{1}{2}(A^2\otimes\mathcal{I}+\mathcal{I}\otimes B^2)\otimes C\otimes\mathcal{I} \ . \tag{3.39}$$

By Lemma 3.4.8 we see that, for sufficiently large Δ, $(H', \mathcal{H}_D)$ defines a $(O(\Delta), \eta, \epsilon)$-gadget for H (see Fig. 3.8(b)).

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

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

$$A = \lambda_0^A|0\rangle\langle 0| + \lambda_1^A|1\rangle\langle 1| \ , \quad B = \lambda_0^B|0\rangle\langle 0| + \lambda_1^B|1\rangle\langle 1| \ , \quad C = \lambda_0^C|0\rangle\langle 0| + \lambda_1^C|1\rangle\langle 1| \ . \tag{3.40}$$

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

$$\begin{aligned} H' = {} & \lambda_0^B(A - \lambda_0^A\mathcal{I})\otimes\mathcal{I}\otimes\mathcal{I}\otimes C \\ & + \lambda_1^B\mathcal{I}\otimes(A - \lambda_0^A\mathcal{I})\otimes\mathcal{I}\otimes C \\ & + \lambda_0^A\mathcal{I}\otimes\mathcal{I}\otimes B\otimes C \ , \end{aligned} \tag{3.41}$$

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

$$P' = (\mathcal{I}\otimes|0\rangle\langle 0|\otimes|0\rangle\langle 0| + |0\rangle\langle 0|\otimes\mathcal{I}\otimes|1\rangle\langle 1|)\otimes\mathcal{I} \ . \tag{3.42}$$

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 $O(\Delta)(\mathcal{I} - P')$ to H'. The caveat is that this gadget has a large η parameter, and hence it does not combine well with other interactions. For instance, in Definition 3.4.2 we might take $P = |0\rangle\langle 0|\otimes\mathcal{I}\otimes\mathcal{I}\otimes\mathcal{I}$, and $U = (\mathbb{F}\otimes|0\rangle\langle 0| + \mathcal{I}\otimes\mathcal{I}\otimes|1\rangle\langle 1|)\otimes\mathcal{I}$, 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. 3.9), and 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.

3.4.3 Gadget combination results

The following results show that gadgets satisfying Definition 3.4.2 or Definition 3.4.5 can be systematically combined as desired. Our techniques and proofs extend prior

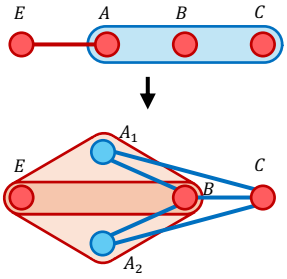


Figure 3.9: The exact 3-to-2 gadget. The (blue) 3-local interaction between A, B, and C is replaced by a series of (blue) 2-local interactions, where the A site has been split into two sites A_1 and A_2. However after this process, the 2-local interaction (red) between A and another qubit E is replaced by two 3-local interactions between E, A_1, B and E, A_2, B. Compare this with Fig. 3.8(b), for which additional interactions on qubit A will remain on qubit A of the gadgetised Hamiltonian without any need for adjustment.

work [135, 72, 171], using the convenient formalism of the direct rotation [172]. The scalings of the parameters η', ϵ' 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 Section 3.6.4.

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

Setup 9. *Let $H \in \mathrm{Herm}(\mathcal{H})$ 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 3.4.2, 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

$$\|H_i'\| \leq J' \ , \quad \|(\mathcal{I} \otimes P_i) H_i' (\mathcal{I} \otimes P_i^{\perp})\| \leq J_O' \ . \tag{3.43}$$

Firstly, we state the main result: that gadgets as in Definition 3.4.2 may be systematically combined to produce new gadgets.

Proposition 3.4.10 (Parallel (η, ϵ)-gadget combination)**.** *Let $H = \sum_i H_i$ be as in Setup 9, and suppose that each $(H_i', \mathcal{A}_i)$ defines a (η, ϵ)-gadget for H_i.*

Define

$$H' = \sum_i H_i' \in \mathrm{Herm}\left(\mathcal{H} \otimes (\otimes_i \mathcal{A}_i)\right) \ . \tag{3.44}$$

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

$$\epsilon' = O(n\epsilon + n\eta J + n\eta^3 J_O' + n\eta^4 J') \ , \quad \eta' = O(n\eta) \ . \tag{3.45}$$

For completeness, we also prove a similar result that (Δ, η, ϵ)-gadgets can be combined to create a new $(\Delta', \eta', \epsilon')$-gadget. It follows from Proposition 3.4.10 that the combination of many (Δ, η, ϵ)-gadgets defines a (η', ϵ')-gadget, however it still remains to show that the projector P' in the sense of Definition 3.4.2 may be taken as a low-energy projector $P_{\leq \Delta'(H')}$.

Proposition 3.4.11 (Parallel (Δ, η, ϵ)-gadget combination)**.** *Let $H = \sum_i H_i$ be as in Setup 9, and suppose that each $(H_i', \mathcal{A}_i)$ defines a (Δ, η, ϵ)-gadget for H_i, where*

$$\Delta \geq \frac{\|H\| + J + N(\epsilon + 2J\eta)}{\frac{1}{4} - 2\eta} = O(nJ) \ , \tag{3.46}$$

and assume that the scaling of η with n is bounded as

$$\eta = o(n^{-\frac{1}{2}}) \ , \tag{3.47}$$

and moreover that, for large J',

$$n\epsilon + n\eta J + n\eta^3 J_O' + n\eta^4 J' = o(J') \ , \quad J' = O(\Delta) \ . \tag{3.48}$$

Define

$$H' = \sum_i H_i' \in \mathrm{Herm}(\mathcal{H} \otimes (\otimes_i \mathcal{A}_i)) \ . \tag{3.49}$$

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

$$\Delta' = \frac{1}{2}\Delta \ , \quad \epsilon' = O(n\epsilon + n\eta J + n\eta^3 J_O' + n^3\eta^4 J') \ , \quad \eta' = O(n\eta) \ . \tag{3.50}$$

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 ϵ' and η' 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 (3.46-3.48) 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 (Δ, η, ϵ) gadgets using Proposition 3.4.11 requires the unappealing conditions of (3.46)-(3.47), which explicitly require the gadget energies to scale with n. In fact, as noted by Bravyi et al.[135], 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 3.4.12 (Ground state energy estimation with (Δ, η, ϵ)-gadgets, generalising Bravyi et al., Theorem 1[135])**.** *Let $H = \sum_i H_i$ be as in Setup 9,and suppose that each $(H_i', \mathcal{A}_i)$ defines a (Δ, η, ϵ)-gadget for H_i.*

Define

$$H' = \sum_i H_i' \in \mathrm{Herm}(\mathcal{H} \otimes (\otimes_i \mathcal{A}_i)) \ . \tag{3.51}$$

Then the ground state energies of H and H' satisfy

$$|\lambda_0(H) - \lambda_0(H')| = O(n\epsilon + n\eta J + n\eta^3 J_O' + n\eta^4 J') \ . \tag{3.52}$$

3.4.4 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 (η, ϵ)-gadgets (Definition 3.4.2), and hence follows even from the relaxed (ζ, ϵ)-gadget property of Definition 3.4.1.

Theorem 3.4.13 (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 $J > 0$,*

$$H = J \bigotimes_{i=1}^{k} Z_i \ . \tag{3.53}$$

Suppose $(H', \mathcal{A})$ is a (η, ϵ)-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 Section 3.6.5) 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 Oliviera et al.[72] and higher-order gadgets[173, 155], 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.[135], and is still unresolved. Using the formalism introduced here, this problem can be precisely stated, and optimisation of Theorem 3.4.13 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 3.4.13 is that it essentially rules out a size-independent (Definition 3.3.1) 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 (ζ, ϵ)-gadget property (Definition 3.4.1) 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 3.4.4, this requires the use of (η, ϵ)-gadgets (Definition 3.4.2) with $\eta = O(n^{-1/2})$, and by Theorem 3.4.13 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[174] gives a method to reduce the exponential or doubly-exponential scaling in perturbative Hamiltonians to polynomial scaling, and Cao et al.[154] present gadgets whose interaction strengths do not grow with accuracy. However, both cases violate size-independence (Definition 3.3.1) in other ways such as polynomial scaling in the number of simulator qubits or instead shrinking the interaction strengths.

3.4.5 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 3.4.13 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 3.3.5 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 3.3.5 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 $H \in \mathrm{Herm}(\mathcal{H})$ 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 $H' \in \mathrm{Herm}(\mathcal{H} \otimes \mathcal{A})$, but now we supplement the natural time evolution with regular projective measurements on the $\mathcal{A}$ system at time intervals of δt. By the quantum Zeno effect [175], 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 3.4.14, provides a formal construction for the measurement-based gadgets described above — see Section 3.6.2 for the proof. Qualitatively, this result tells us that if we evolve $|\psi\rangle \otimes |0\rangle$ for time δt 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 3.4.14. *For a Hilbert space $\mathcal{H}$ and an ancillary qubit $\mathcal{A} = \mathbb{C}^2$, let $H' \in \mathrm{Herm}(\mathcal{H} \otimes \mathcal{A})$ be a Hamiltonian given by*

$$H' = H_{\mathcal{I}} \otimes \mathcal{I} + H_X \otimes X + H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1| \ , \tag{3.54}$$

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

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

$$e^{-i\delta t H'}(|\psi\rangle \otimes |0\rangle) = \left(e^{-i\delta t H}|\psi\rangle + O((\delta t)^2)\right) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle \ , \tag{3.55}$$

where

$$H = H_{\mathcal{I}} - \omega^{-2} H_X H_{|1\rangle\langle 1|} H_X \ . \tag{3.56}$$

This provides a new 3-to-2-local gadget for Pauli strings. For example, we can set $H_{\mathcal{I}} = -Z_1$, $H_X = \sqrt{\frac{\omega}{2}}(Z_2 + Z_3)$, $H_{|1\rangle\langle 1|} = -\omega Z_1$; 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 $H_{\mathcal{I}} = -A_a$, $H_X = \sqrt{\frac{\omega}{2}}(B_b + C_c)$, $H_{|1\rangle\langle 1|} = -\omega A_a$ 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 3.4.14 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 3.4.15 provides the necessary result for this, by verifying that the conclusions of Proposition 3.4.14 also hold when an additional term $H_{\text{else}} \in \text{Herm}(\mathcal{H})$ is added to both the target and simulator Hamiltonian.

Proposition 3.4.15. *Let* $H_{else} = \sum_i h_i$ *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*

$$H' = H_{\mathcal{I}} \otimes \mathcal{I} + H_X \otimes X + H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1| \ , \tag{3.57}$$

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

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

$$e^{-i\delta t(H' + H_{else} \otimes \mathcal{I})}(|\psi\rangle \otimes |0\rangle) = \left(e^{-i\delta t(H + H_{else})}|\psi\rangle + O((\delta t)^2)\right) \otimes |0\rangle + O((\delta t)^{3/2}) \otimes |1\rangle \ , \tag{3.58}$$

where

$$H = H_{\mathcal{I}} - \omega^{-2} H_X H_{|1\rangle\langle 1|} H_X \ . \tag{3.59}$$

The significance of Proposition 3.4.15 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 $e^{-i\delta t(H + H_{\text{else}})} \approx e^{-i\delta t H} e^{-i\delta t H_{\text{else}}}$.

3.5 Discussion

Given the result of Proposition 3.4.15, we can now describe how the measurement gadget construction fits into our framework of analogue quantum simulation described in Definition 3.3.5.

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 3.4.15, 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 $H' \in \mathrm{Herm}(\mathcal{H}')$ be given by

$$H' = -Z_1 \otimes \mathcal{I} + \sqrt{\frac{\omega}{2}}(Z_2 + Z_3) \otimes X - \omega Z_1 \otimes |1\rangle\langle 1| \ , \tag{3.60}$$

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

$$\mathcal{E}_{\mathrm{state}}(\rho) = \rho \otimes |0\rangle\langle 0| \ , \quad \mathcal{E}_{\mathrm{obs}}(O) = O \otimes \mathcal{I} \ , \tag{3.61}$$

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

$$E_{\delta t}(\rho') = e^{-i\delta t(H' + H_{\mathrm{else}} \otimes \mathcal{I})} \rho' e^{i\delta t(H' + H_{\mathrm{else}} \otimes \mathcal{I})} \ , \tag{3.62}$$

$$M(\rho') = \mathrm{tr}_{\mathcal{A}}[\rho'(\mathcal{I} \otimes |0\rangle\langle 0|)] \otimes |0\rangle\langle 0| + \mathrm{tr}_{\mathcal{A}}[\rho'(\mathcal{I} \otimes |1\rangle\langle 1|] \otimes |1\rangle\langle 1| \ , \tag{3.63}$$

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

$$T_t = (M \circ E_{\delta t}) \circ (M \circ E_{\delta t}) \circ \cdots \circ (M \circ E_{\delta t}) \ , \tag{3.64}$$

containing $\lfloor t/\delta t \rfloor$ copies of $(M \circ E_{\delta t})$. This evolution is described by Fig. 3.4. The content of Proposition 3.4.15 tells us that

$$(T_t \circ \mathcal{E}_{\mathrm{state}})(\rho) = \left(e^{-itH} \rho e^{itH} + O(t\delta t)\right) \otimes |0\rangle\langle 0| + O(t(\delta t)^2) \otimes |1\rangle\langle 1| \ , \tag{3.65}$$

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

$$\mathrm{tr}[\mathcal{E}_{\mathrm{obs}}(O)(T_t \circ \mathcal{E}_{\mathrm{state}})(\rho)] = \mathrm{tr}[O e^{-itH} \rho e^{itH}] + O(t\delta t) \ . \tag{3.66}$$

The channels T_t therefore simulate H (in the sense of Definition 3.3.5) with respect to any states Ω_{state} and normalised observables Ω_{obs}, up to accuracy $\epsilon > 0$ and maximum time $t_{\max}$, provided that one chooses $\delta t = O(\epsilon t_{\max}^{-1})$. Therefore we require interaction strengths and measurement frequency which scale as $J = O(\epsilon^{-1} t_{\max})$ — 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 (η, ϵ)-gadget in the sense of Definition 3.4.2, 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 3.4.13, this must involve energy scalings of $J = \Omega(\epsilon^{-1} n)$ (and even without Theorem 3.4.13, a low-energy (Δ, η, ϵ)-gadget as in Definition 3.4.5 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 $n = O(\log^d(1/\epsilon) + t_{\max}^d)$ (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_{gad} such gadgets were used in parallel, we expect (though do not calculate here) that an additional overhead of at least $\delta t = O\big((t_{\max} N_{\text{gad}})^{-1/2}\big)$ 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 3.3.5. 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 3.4.2 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.

3.6 Appendix

3.6.1 Local encodings

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

Proof of Proposition 3.3.7. 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

$$\begin{aligned} X'_k \mathcal{E}_{\text{state}}(\rho)(X'_k)^\dagger &= X'_k \operatorname{tr}_E[U(\rho \otimes |0\rangle\langle 0|_F)U^\dagger](X'_k)^\dagger \\ &= \operatorname{tr}_E[U X_k(\rho \otimes |0\rangle\langle 0|_F)X_k^\dagger U^\dagger] \,, \end{aligned} \tag{3.67}$$

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

∎

Proof of Proposition 3.3.9. To see this, we use the definition of local state encodings and write

$$\mathcal{E}^*_{obs}(\rho') = \mathrm{tr}_G[W(\rho' \otimes |0\rangle\langle 0|_E)W^\dagger] \ , \tag{3.68}$$

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

$$\begin{aligned} \mathrm{tr}[\mathcal{E}_{\mathrm{obs}}(O)\rho'] &= \mathrm{tr}[O\mathcal{E}^*_{\mathrm{obs}}(\rho')] \\ &= \mathrm{tr}[(O \otimes \mathcal{I}_G)W(\rho' \otimes |0\rangle\langle 0|_E)W^\dagger] \\ &= \mathrm{tr}[(\mathcal{I}_{\mathcal{H}'} \otimes \langle 0|_E)W^\dagger(O \otimes \mathcal{I}_G)W(\mathcal{I}_{\mathcal{H}'} \otimes |0\rangle_E)\rho'] \ . \end{aligned} \tag{3.69}$$

Assuming O is local, then $W^\dagger(O \otimes \mathcal{I}_G)W$ 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

$$W^\dagger(O \otimes \mathcal{I}_G)W = O' \otimes \mathcal{I}_{A^c} \ , \tag{3.70}$$

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

$$\mathrm{tr}[\mathcal{E}_{\mathrm{obs}}(O)\rho'] = \mathrm{tr}[(\mathcal{I}_A \otimes \langle 0|_E)O'(\mathcal{I}_A \otimes |0\rangle_E)\rho'_A] \ , \tag{3.71}$$

which can be estimated via a POVM on A. ∎

3.6.2 Qutrit-to-qubit energy scaling

The idea for the proof of Proposition 3.3.10 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 (Δ, η, ϵ) 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 Δ) subspace.

In the proof below, we start with the encoded ground state ρ_0, and construct a state ρ_1 which differs only from ρ_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[176].

Lemma 3.6.1 (Weyl's Perturbation Theorem)**.** *Let $A, B \in \mathrm{Herm}(\mathcal{H})$ be Hermitian matrices, with spectra $\lambda_0 \leq \lambda_1 \leq \dots$ and $\mu_0 \leq \mu_1 \leq \dots$ respectively. Then*

$$\max_j |\lambda_j - \mu_j| \leq \|A - B\| \ . \tag{3.72}$$

Proof of Proposition 3.3.10. 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.[56], we have two encodings $\mathcal{E}_{\mathrm{obs}}$ and $\tilde{\mathcal{E}}_{\mathrm{obs}}$ of the form (using that H is real to set $Q = 0$ without loss of generality)

$$\mathcal{E}_{\mathrm{obs}}(M) = V(M \otimes P)V^\dagger \ , \quad \tilde{\mathcal{E}}_{\mathrm{obs}}(M) = \tilde{V}(M \otimes P)\tilde{V}^\dagger \ , \tag{3.73}$$

where P is a projector on the ancillary space $\mathcal{A}$, and $V, \tilde{V}$ are both isometries $\mathcal{H} \otimes \mathcal{A} \to \mathcal{H}'$. 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 $V_i : \mathcal{H}_i \otimes \mathcal{A}_i \to \mathcal{H}'_i$. 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
$$\tilde{\mathcal{E}}_{\text{obs}}(\mathcal{I}) = \tilde{V}(\mathcal{I} \otimes P)\tilde{V}^\dagger = P_{\leq \Delta(H'_n)} \ , \tag{3.74}$$
where $P_{\leq \Delta(H'_n)}$ is the low-energy (below Δ) projector for H'_n, and
$$\| P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} - \tilde{\mathcal{E}}_{\text{obs}}(H_n) \| \leq \epsilon \ . \tag{3.75}$$

- $\mathcal{E}_{\text{obs}}$ and $\tilde{\mathcal{E}}_{\text{obs}}$ are close, in the sense that
$$\| V - \tilde{V} \| \leq \eta \ . \tag{3.76}$$

Now we define a state $\tau \in \text{span}(P)$ and define a state encoding (in the sense of Cubitt et al.[56])
$$\tilde{\mathcal{E}}_{\text{state}}(\rho) = \tilde{V}(\rho \otimes \tau)\tilde{V}^\dagger \ . \tag{3.77}$$
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
$$P_{\leq \Delta(H'_n)} \rho_0 = \rho_0 \ , \quad \tilde{\mathcal{E}}_{\text{obs}}(H_n)\rho_0 = 0 \ . \tag{3.78}$$
Hence we can bound the energy of ρ_0 under H'_n by
$$\begin{aligned} \text{tr}[H'_n \rho_0] &= \text{tr}[P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} \rho_0] \\ &= \text{tr}[(P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} - \tilde{\mathcal{E}}_{\text{obs}}(H_n))\rho_0] \\ &\leq \| P_{\leq \Delta(H'_n)} H'_n P_{\leq \Delta(H'_n)} - \tilde{\mathcal{E}}_{\text{obs}}(H_n) \| \\ &\leq \epsilon \ . \end{aligned} \tag{3.79}$$
Now without loss of generality we assume that $m_1 = \min_i m_i$. Notice that $V_1 : \mathcal{H}_1 \otimes \mathcal{A}_1 \to \mathcal{H}'_1$ cannot be surjective, since
$$\dim(\mathcal{H}_1 \otimes \mathcal{A}_1) = 3 \dim \mathcal{A}_1 \neq 2^{m_1} \ . \tag{3.80}$$
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
$$\rho_1 = \psi \otimes \text{tr}_1[\rho_0] \in \text{Lin}(\mathcal{H}') \ , \tag{3.81}$$
Where 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
$$\begin{aligned} \text{tr}[P_{\leq \Delta(H'_n)} \rho_1] &= \text{tr}[(\mathcal{I} \otimes P)\tilde{V}^\dagger \rho_1 \tilde{V}] \\ &\leq \text{tr}[\tilde{V}^\dagger \rho_1 \tilde{V}] \\ &= \text{tr}[(\tilde{V} - V)^\dagger \rho_1 (\tilde{V} - V)] \\ &\leq \| (\tilde{V} - V)(\tilde{V} - V)^\dagger \| \\ &\leq \eta^2 \ , \end{aligned} \tag{3.82}$$

from which we deduce that

$$\operatorname{tr}[H'_n\rho_1] \geq \Delta \operatorname{tr}[(\mathcal{I} - P_{\leq\Delta(H'_n)})\rho_1] - \epsilon \operatorname{tr}[P_{\leq\Delta(H'_n)}\rho_1] \geq \Delta(1-\eta^2) - \epsilon\eta^2 \ , \tag{3.83}$$

using that the smallest eigenvalue of H'_n is at least $-\epsilon$, by (3.75) and Lemma 3.6.1. Therefore, using (3.79),

$$\operatorname{tr}[H'_n(\rho_1 - \rho_0)] \geq \Delta(1-\eta^2) - \epsilon(1+\eta^2) \ . \tag{3.84}$$

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

$$\operatorname{tr}[H'_n(\rho_1 - \rho_0)] = \sum_{j=1}^{K} \operatorname{tr}[h'_j(\rho_1 - \rho_0)] \ . \tag{3.85}$$

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

$$\begin{aligned} \operatorname{tr}[h'_j(\rho_1 - \rho_0)] &= \operatorname{tr}[(\mathcal{I}_1 \otimes \tilde{h}_j)(\psi \otimes \operatorname{tr}_1 \rho_0 - \rho_0)] \\ &= \operatorname{tr}_1[\psi] \operatorname{tr}_{2,3,\ldots}[\tilde{h}_j \operatorname{tr}_1[\rho_0]] - \operatorname{tr}[(\mathcal{I} \otimes \tilde{h}_j)\rho_0] \\ &= 0 \ . \end{aligned} \tag{3.86}$$

Hence the only non-zero contributions to (3.85) come from j in the set

$$I_1 = \{1 \leq j \leq K \mid h'_j \text{ acts non-trivially on } \mathcal{H}'_1\} \ . \tag{3.87}$$

So (3.85) can be bounded by

$$\operatorname{tr}[H'_n(\rho_1 - \rho_0)] = \sum_{j\in I_1} \operatorname{tr}[h'_j(\rho_1 - \rho_0)] \leq 2|I_1| \max_{j\in I_1} \|h'_j\| \ , \tag{3.88}$$

using the Hölder inequality for Schatten p-norms.

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

$$\|\tilde{\mathcal{E}}_{\text{obs}}(H_n)\| = \|H_n\| = n \ , \tag{3.89}$$

so by (3.75) and Lemma 3.6.1

$$\|P_{\leq\Delta(H'_n)} H'_n P_{\leq\Delta(H'_n)}\| \geq n - \epsilon \ . \tag{3.90}$$

Hence, by the definition of $P_{\leq\Delta(H'_n)}$, we must have $\Delta > n - \epsilon$. Combining this fact with (3.84) and (3.88), we deduce that

$$\max_{j\in I_1} \|h'_j\| > \frac{1}{2|I_1|}\big((n-\epsilon)(1-\eta^2) - \epsilon(1+\eta^2)\big) \ . \tag{3.91}$$

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

$$\max_{j\in I_1} \|h'_j\| \geq \Omega\big(n^{1-\alpha}(1-\eta^2)\big) \ . \tag{3.92}$$

■

Proof of Lemma 3.3.11. 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

$$|B_R(y)| = O(n^{1-\alpha} + n^{1-1/D}) = O(n^{1-\min\{\alpha,1/D\}}) \ . \tag{3.93}$$

■

3.6.3 Gadget characterisation

In this section, we give the proofs of Theorems 3.4.3 and 3.4.4, 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.[172]. We summarise the basic definitions and properties here without proof.

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 $U_{\psi\to\phi}$ which rotates only within the subspace spanned by $|\psi\rangle$ and $|\phi\rangle$. Defining the reflections $R_\psi = \mathcal{I} - 2|\psi\rangle\langle\psi|$ and $R_\phi = \mathcal{I} - 2|\phi\rangle\langle\phi|$, we can write

$$U_{\psi\to\phi} = \sqrt{R_\phi R_\psi} \ , \tag{3.94}$$

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

$$R_\mathcal{P} = \mathcal{I} - 2P \ , \quad R_\mathcal{Q} = \mathcal{I} - 2Q \ , \tag{3.95}$$

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

$$U_{\mathcal{P}\to\mathcal{Q}} = \sqrt{R_\mathcal{Q} R_\mathcal{P}} \ , \tag{3.96}$$

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 $U_{\mathcal{P}\to\mathcal{Q}}$ satisfies

$$U_{\mathcal{P}\to\mathcal{Q}} P U_{\mathcal{P}\to\mathcal{Q}}^\dagger = Q \ . \tag{3.97}$$

Moreover, as described by Bravyi et al. [172], the direct rotation may be written in terms of its generator: an anti-Hermitian operator $S = -S^\dagger$ which can be chosen so that $U_{\mathcal{P}\to\mathcal{Q}} = e^S$, with $\|S\| < \pi/2$ and which is off-diagonal with respect to both P and Q:

$$PSP = (\mathcal{I}-P)S(\mathcal{I}-P) = QSQ = (\mathcal{I}-Q)S(\mathcal{I}-Q) = 0 \ . \tag{3.98}$$

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

$$\|U_{\mathcal{P}\to\mathcal{Q}} - \mathcal{I}\,\| = \max_j |2\sin(\theta_j/2)| \ , \quad \|S\| = \max_j |\theta_j| \ , \tag{3.99}$$

and hence

$$\|S\| \leq \frac{\pi}{2\sqrt{2}} \|U_{\mathcal{P}\to\mathcal{Q}} - \mathcal{I}\,\| \ . \tag{3.100}$$

The gadget property from gadgets

Proof of Theorem 3.4.3. This follows directly from the definition, since for any $H_{\mathrm{else}} \in \mathrm{Herm}(\mathcal{H})$, we have

$$\begin{aligned} &\|P'(H' + H_{\mathrm{else}} \otimes \mathcal{I})P' - U\big((H + H_{\mathrm{else}}) \otimes P\big)U^\dagger\| \\ &\quad \leq \|P'H'P' - U(H \otimes P)U^\dagger\| + \|P'(H_{\mathrm{else}} \otimes \mathcal{I})P' - U(H_{\mathrm{else}} \otimes P)U^\dagger\| \ . \end{aligned} \tag{3.101}$$

The first term is bounded by ϵ by definition, and the second term can be bounded using

$$\begin{aligned} &\|P'(H_{\mathrm{else}} \otimes \mathcal{I})P' - U(H_{\mathrm{else}} \otimes P)U^\dagger\| \\ &\quad = \|(\mathcal{I}\otimes P)U^\dagger(H_{\mathrm{else}} \otimes \mathcal{I})U(\mathcal{I}\otimes P) - H_{\mathrm{else}} \otimes P\| \\ &\quad \leq 2\eta\|H_{\mathrm{else}}\| \ . \end{aligned} \tag{3.102}$$

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

Gadgets from the gadget property

The proof of Theorem 3.4.4 requires Lemma 3.6.3, 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 3.6.3. *Let $P, Q \in \mathrm{Proj}(\mathcal{H})$ be projectors, and define*

$$f(P, Q) := \min_{\tilde{P}\in\mathrm{Proj}(\mathcal{H})} \|PQP - \tilde{P}\| \ . \tag{3.103}$$

Then

$$\|[P, Q]\| = \sqrt{f(P, Q) - f(P, Q)^2} \ . \tag{3.104}$$

Proof of Lemma 3.6.3. Write P and Q in the block-diagonal basis of P, so that

$$P = \begin{pmatrix} \mathcal{I} & 0 \\ 0 & 0 \end{pmatrix} , \quad Q = \begin{pmatrix} A & B \\ B^\dagger & C \end{pmatrix} , \tag{3.105}$$

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

$$PQP = \begin{pmatrix} A & 0 \\ 0 & 0 \end{pmatrix} . \tag{3.106}$$

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

$$f(P, Q) = \max_j \big(\min\{|\lambda_j|, |1 - \lambda_j|\} \big) . \tag{3.107}$$

To see why (3.107) 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 3.6.1.

Now we can compute

$$-[P, Q]^2 = \begin{pmatrix} BB^\dagger & 0 \\ 0 & B^\dagger B \end{pmatrix} , \tag{3.108}$$

hence

$$\|[P, Q]\|^2 = \| - [P, Q]^2 \| = \|BB^\dagger\| = \|A(\mathcal{I} - A)\| = \max_j |\lambda_j||1 - \lambda_j| . \tag{3.109}$$

Note that the maximising j in (3.107) and (3.109) must be the same (the functions $\min\{|\lambda|, |1 - \lambda|\}$ and $|\lambda||1 - \lambda|$ are both maximised by the λ_j closest to $1/2$), hence we can deduce

$$\|[P, Q]\|^2 = f(P, Q)\big(1 - f(P, Q)\big) , \tag{3.110}$$

which gives the result. ∎

In order to obtain the correct unitary U in the gadget definition, the proof of Theorem 3.4.4 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 3.4.11.

Lemma 3.6.4 (Davis-Kahan $\sin\theta$ theorem [177])**.** *Let $A, B \in \mathrm{Herm}(\mathcal{H})$, and take $P_A, P_B \in \mathrm{Proj}(\mathcal{H})$ projectors of equal rank which block-diagonalise A and B respectively, so that*

$$A = P_A A P_A + P_A^\perp A P_A^\perp , \quad B = P_B B P_B + P_B^\perp B P_B^\perp . \tag{3.111}$$

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

$$\mathrm{spec}\big(A|_{P_A\mathcal{H}}\big) \subset [\alpha, \beta] , \quad \mathrm{spec}\big(B|_{P_B^\perp\mathcal{H}}\big) \subset \mathbb{R} \setminus (\alpha - \lambda_{\mathrm{gap}}, \beta + \lambda_{\mathrm{gap}}) . \tag{3.112}$$

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

$$\|U - \mathcal{I}\| \leq \frac{\sqrt{2}}{\lambda_{\mathrm{gap}}} \|(B - A)P_A\| . \tag{3.113}$$

Proof of Lemma 3.6.4. The statement of Davis et al.[177] is phrased in terms of a matrix $\Theta_0 = \mathrm{diag}(\theta_1, \theta_2, \ldots, \theta_n)$, where the eigenvalues (possibly excluding some 1's) of U are given by $e^{i\theta_j}$ and $\pi/2 \geq \theta_1 \geq \theta_2 \geq \cdots \geq \theta_n$. Specifically, the authors give the following result:

$$\lambda_{\mathrm{gap}} \| \sin \Theta_0 \| \leq \| (B - A) P_A \| \ . \tag{3.114}$$

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

$$\| U - \mathcal{I} \| = |2 \sin(\theta_1/2)| \leq \sqrt{2} |\sin \theta_1| = \sqrt{2} \| \sin \Theta_0 \| \ . \tag{3.115}$$

■

The following lemmas from Bravyi et al.[135] 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

$$e^S H e^{-S} = H + [S, H] + \frac{1}{2!}[S, [S, H]] + \frac{1}{3!}[S, [S, [S, H]]] + \ldots \ . \tag{3.116}$$

Lemma 3.6.5 (Bravyi et al., Lemma 1[135])**.** *Let S be an anti-Hermitian operator. Define a superoperator ad_S such that $\mathrm{ad}_S(X) = [S, X]$, and let ad_S^k be the k-fold composition of ad_S, with $\mathrm{ad}_S^0(X) = X$. For any operator H define $r_0(H) = \| e^S H e^{-S} \| = \| H \|$, $r_1(H) = \| e^S H e^{-S} - H \|$, and*

$$r_k(H) = \| e^S H e^{-S} - \sum_{p=0}^{k-1} \frac{1}{p!} \mathrm{ad}_S^p(H) \| \ , \quad k \geq 2 \ . \tag{3.117}$$

Then for all $k \geq 0$ one has

$$r_k(H) \leq \frac{1}{k!} \| \mathrm{ad}_S^k(H) \| \ . \tag{3.118}$$

Lemma 3.6.6 (Bravyi et al., Lemma 2[135])**.** *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 $k = O(1)$,*

$$\| \mathrm{ad}_S^k(H) \| = O(n J_S^k J_H) \ . \tag{3.119}$$

These lemmas provide us with the necessary tools to prove Theorem 3.4.4.

Proof of Theorem 3.4.4. The idea of the proof is as follows.

By the gadget property, we have

$$\| P'(H' + H_{\mathrm{else}} \otimes \mathcal{I}) P' - \tilde{U}_{H_{\mathrm{else}}} \big((H + H_{\mathrm{else}}) \otimes \tilde{P} \big) \tilde{U}_{H_{\mathrm{else}}}^\dagger \| \leq \epsilon + \zeta \| H_{\mathrm{else}} \| \ , \tag{3.120}$$

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

1. First we consider (3.120) the case where H_{else} dominates the expression, and argue that that P' is "almost" a projector $\mathcal{I} \otimes P$ on $\mathcal{A}$.

2. Next, by setting $H_{\text{else}} = 0$ in (3.120), we observe that $P'H'P'$ has approximately the same spectrum as $H \otimes P$.

3. By setting $H_{\text{else}} = -H$ in (3.120), we argue that $P'H'P' \approx H \otimes P$.

4. Using steps 2 and 3, and Lemma 3.6.4, 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 (3.120) yields

$$\|P'(H_{\text{else}} \otimes \mathcal{I})P' - \tilde{U}_{\lambda H_{\text{else}}}(H_{\text{else}} \otimes \tilde{P})\tilde{U}^\dagger_{\lambda H_{\text{else}}}\| \leq \zeta + O(\lambda^{-1}) \ , \tag{3.121}$$

for large λ, which in particular, by Lemma 3.6.1, implies that $P'(H_{\text{else}} \otimes \mathcal{I})P'$ has the same spectrum as $H_{\text{else}} \otimes \tilde{P}$, up to error ζ. (That is, the kth smallest eigenvalue, counted with multiplicity, of $P'(H_{\text{else}} \otimes \mathcal{I})P'$, differs from that of $H_{\text{else}} \otimes \tilde{P}$ by an absolute error of at most ζ.)

In particular, if $H_{\text{else}} = Q \in \text{Proj}(\mathcal{H})$ is a projection, then so is $P'(Q \otimes \mathcal{I})P'$ (up to spectral error ζ). Hence by Lemma 3.6.3, we have

$$\|[P', Q \otimes \mathcal{I}]\| \leq \sqrt{\zeta} \ . \tag{3.122}$$

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_{else} as a linear combination of at most K projections, we have

$$\|[P', H_{\text{else}} \otimes \mathcal{I}]\| \leq \pi K \sqrt{\zeta} \ , \tag{3.123}$$

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

Therefore, for any $V \in \text{U}(\mathcal{H})$, we can write $V = e^{iH_{\text{else}}}$ for some H_{else} as above, and then by Lemma 3.6.5,

$$\|(V \otimes \mathcal{I})P'(V^\dagger \otimes \mathcal{I}) - P'\| \leq \pi K \sqrt{\zeta} \ . \tag{3.124}$$

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

$$\left\| \int \text{d}V (V \otimes \mathcal{I})P'(V^\dagger \otimes \mathcal{I}) - P' \right\| \leq \pi K \sqrt{\zeta} \ , \tag{3.125}$$

but

$$\int \text{d}V (V \otimes \mathcal{I})P'(V^\dagger \otimes \mathcal{I}) = K^{-1} \, \mathcal{I}_{\mathcal{H}} \otimes \text{tr}_{\mathcal{H}}[P'] \ , \tag{3.126}$$

hence

$$\|P' - \mathcal{I}_{\mathcal{H}} \otimes K^{-1} \, \text{tr}_{\mathcal{H}}[P']\| \leq \pi K \sqrt{\zeta} \ . \tag{3.127}$$

In particular, by Lemma 3.6.1 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 ζ 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

$$\|P' - \mathcal{I} \otimes P\| \leq 2\pi K \sqrt{\zeta} \ . \tag{3.128}$$

Now we can apply Lemma 3.6.4, using $A = P_A = P'$, $B = P_B = \mathcal{I} \otimes P$, and $\lambda_{\text{gap}} = 1$. Then for ζ small enough, the direct rotation W from $\mathcal{I} \otimes P$ to P' satisfies

$$P' = W(\mathcal{I} \otimes P)W^\dagger \ , \quad \|W - \mathcal{I}\| \leq 2\sqrt{2}\pi K \sqrt{\zeta} \ , \tag{3.129}$$

which completes step 1. Without loss of generality we can now adjust the $\tilde{U}_{H_{\text{else}}}$ so that $\tilde{P} = P$, since (3.121) holds for all H_{else} and therefore $\operatorname{rank} P' = \operatorname{rank} \mathcal{I} \otimes P = \operatorname{rank} \mathcal{I} \otimes \tilde{P}$ for small enough ζ. Hence for all $H_{\text{else}} \in \text{Herm}(\mathcal{H})$ we have

$$\|P'(H' + H_{\text{else}} \otimes \mathcal{I})P' - \tilde{U}_{H_{\text{else}}}\big((H + H_{\text{else}}) \otimes P\big)\tilde{U}_{H_{\text{else}}}^\dagger\| \leq \epsilon + \zeta\|H_{\text{else}}\| \ . \tag{3.130}$$

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

$$\|P'H'P' - \tilde{U}(0)(H \otimes P)\tilde{U}^\dagger(0)\| \leq \epsilon \ , \tag{3.131}$$

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

$$\|P'H'P' - P'(H \otimes \mathcal{I})P'\| \leq \epsilon + \zeta\|H\| \ . \tag{3.132}$$

Moreover, by (3.129) we can bound

$$\begin{aligned}
\|P'(H \otimes \mathcal{I})P' - H \otimes P\| &= \|W(\mathcal{I} \otimes P)W^\dagger(H \otimes \mathcal{I})W(\mathcal{I} \otimes P)W^\dagger - H \otimes P\| \\
&= \|(W - \mathcal{I})(\mathcal{I} \otimes P)W^\dagger(H \otimes \mathcal{I})W(\mathcal{I} \otimes P)W^\dagger \\
&\quad + (\mathcal{I} \otimes P)(W^\dagger - \mathcal{I})(H \otimes \mathcal{I})W(\mathcal{I} \otimes P)W^\dagger \\
&\quad + (H \otimes P)(W - \mathcal{I})(\mathcal{I} \otimes P)W^\dagger \\
&\quad + (H \otimes P)(W^\dagger - \mathcal{I})\| \\
&\leq 4\|H\| \cdot \|W - \mathcal{I}\| \\
&\leq 8\sqrt{2}\pi K\sqrt{\zeta}\|H\| \ .
\end{aligned} \tag{3.133}$$

Combining (3.132) and (3.133), we complete step 3:

$$\|P'H'P' - H \otimes P\| \leq \epsilon + (\zeta + 8\sqrt{2}\pi K\sqrt{\zeta})\|H\| := \delta \ . \tag{3.134}$$

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 $0 < \lambda_2 < \cdots < \lambda_{M+1}$ 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 ϵ which does not affect the conclusions of this theorem. To see this, note that for $\mu \in \mathbb{R}$,

$$\begin{aligned}
\|P'(H' + \mu\mathcal{I} + H_{\text{else}} \otimes \mathcal{I})P' - \tilde{U}_{H_{\text{else}}}\big((H + \mu\mathcal{I} + H_{\text{else}}) \otimes \tilde{P}\big)\tilde{U}_{H_{\text{else}}}^\dagger\| &\leq \epsilon + \zeta\|\mu\mathcal{I} + H_{\text{else}}\| \\
&\leq (\epsilon + \mu\zeta) + \zeta\|H_{\text{else}}\| \ ,
\end{aligned} \tag{3.135}$$

using (3.120) and absorbing $\mu\mathcal{I}$ into H_{else}. Hence $H' \mapsto H' + \mu\mathcal{I}$ and $H \mapsto H + \mu\mathcal{I}$ also satisfy the assumptions of the theorem, up to replacing $\epsilon \mapsto \epsilon + \mu\zeta$. Moreover, if we can show that $H' + \mu\mathcal{I}$ is an (ϵ, η)-gadget for $H + \mu\mathcal{I}$, then applying the gadget definition shows that

$$\begin{aligned} \|P'H'P' - U(H \otimes P)U^\dagger\| &= \|P'(H' + \mu\mathcal{I})P' - U\big((H + \mu\mathcal{I}) \otimes P\big)U^\dagger\| \\ &\leq \epsilon \ , \end{aligned} \tag{3.136}$$

hence H' is an (ϵ, η)-gadget for H.

We define

$$\lambda_{\text{gap}} = \min_{j \neq k} |\lambda_j - \lambda_k| \ , \tag{3.137}$$

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

We also diagonalise $P'H'P'$ — see Fig. 3.10. By (3.131) and Weyl's inequality we can write the eigenvalues as $\{\mu_k^{(i_k)}\}$ such that

$$|\mu_k^{(i_k)} - \lambda_k| \leq \epsilon \ , \quad \text{for all } i_k, \text{ and for all } k. \tag{3.138}$$

Let $\mathcal{P}'_k$ be the eigenspace of $P'H'P'$ corresponding to the eigenvalues $\{\mu_k^{(i_k)}\}_{i_k}$, which by (3.131) satisfies $\dim \mathcal{P}'_k = \dim \mathcal{P}_k$ for ϵ sufficiently small. Note that for $j \neq k$ we have

$$|\mu_j^{(i_j)} - \lambda_k| \geq \lambda_{\text{gap}} - \epsilon \ . \tag{3.139}$$

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 $i = 1, \ldots, k$. Moreover the 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 3.6.4; we will see that the direct rotations are well-defined for sufficiently small ζ and ϵ.

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

$$\omega_k = \frac{\delta}{2\|H\|}\left(\left[1 + \frac{2\sqrt{2}\|H\|}{\lambda_{\text{gap}} - \epsilon}\right]^k - 1\right) . \tag{3.140}$$

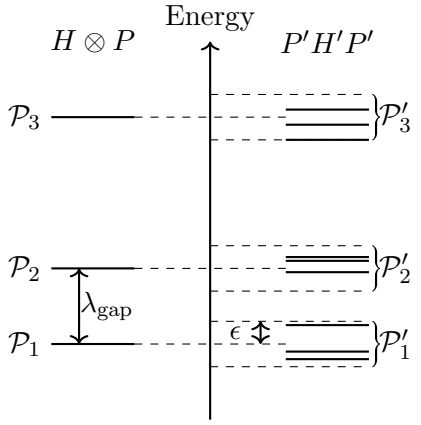


Figure 3.10: The degenerate eigenvalues of $H \otimes P$ are approximated by eigenvalues of $P'H'P'$ up to error ϵ, corresponding to eigenspaces $\mathcal{P}'_i$. We construct a unitary operator which rotates the eigenspaces $\mathcal{P}_i$ onto the $\mathcal{P}'_i$

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 $W^{(0)} = \mathcal{I}$.

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

$$\|P'H'P' - H^{(k-1)}\| \leq \delta + 2\omega_{k-1}\|H\| \ , \tag{3.141}$$

using (3.134) and the bound $\|W^{(k-1)} - \mathcal{I}\| \leq \omega_{k-1}$. Hence, applying Lemma 3.6.4 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

$$\begin{aligned}\|V^{(k)} - \mathcal{I}\| &\leq \frac{\sqrt{2}}{\lambda_{\text{gap}} - \epsilon}\|(P'H'P')|_{\oplus_{j\geq k}\mathcal{P}'_j} - (H^{(k-1)})|_{\oplus_{j\geq k}\mathcal{P}'_j}\| \\ &\leq \frac{\sqrt{2}}{\lambda_{\text{gap}} - \epsilon}(\delta + 2\omega_{k-1}\|H\|) \ . \end{aligned} \tag{3.142}$$

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 $|\psi\rangle \in W^{(k-1)}\mathcal{P}(W^{(k-1)})^\dagger \cap (\mathcal{P}'_k)^\perp$. Then, by the definition of these subspaces,

$$\langle\psi|P'H'P'|\psi\rangle \geq \lambda_{k+1} - \epsilon \geq \lambda_k + (\lambda_{\text{gap}} - \epsilon) \ , \quad \langle\psi|H^{(k-1)}|\psi\rangle = \lambda_k \ , \tag{3.143}$$

which would imply that

$$\|P'H'P' - H^{(k-1)}\| \geq \lambda_{\text{gap}} - \epsilon \ . \tag{3.144}$$

By (3.141), this is prohibited for sufficiently small ζ and ϵ — so we can safely assume that the direct rotation is well-defined.

Now we can let

$$W^{(k)} = (\mathcal{I}_{\oplus_{j<k}\mathcal{P}'_j} \oplus V^{(k)})W^{(k-1)} \ , \tag{3.145}$$

and

$$H^{(k)} = W^{(k)}H^{(0)}(W^{(k)})^\dagger \ , \tag{3.146}$$

which satisfies

$$\|W^{(k)} - \mathcal{I}\| \leq \omega_{k-1} + \frac{\sqrt{2}}{\lambda_{\text{gap}} - \epsilon}(\delta + 2\omega_{k-1}\|H\|) = \omega_k \ . \tag{3.147}$$

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

$$H^{(M+1)} = W^{(M+1)}(H \otimes P)(W^{(M+1)})^\dagger \ , \tag{3.148}$$

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

$$\|P'H'P' - W^{(M+1)}(H \otimes P)(W^{(M+1)})^\dagger\| \leq \epsilon \ . \tag{3.149}$$

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 $(\mathcal{I} \otimes P)$, and $\text{rank}(H \otimes P) = \text{rank}(\mathcal{I} \otimes P)$. By (3.129) and (3.131) we know that $\text{rank}(P') = \text{rank}(\mathcal{I} \otimes P)$ and $\text{rank}(P'H'P') = \text{rank}(H \otimes P)$ for sufficiently small ζ and ϵ. So

$$\text{rank}(P') = \text{rank}(\mathcal{I} \otimes P) = \text{rank}(H \otimes P) = \text{rank}(P'H'P') \ , \tag{3.150}$$

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

$$P' = W^{(M+1)}(\mathcal{I} \otimes P)(W^{(M+1)})^\dagger \ . \tag{3.151}$$

We have therefore shown that $(H', \mathcal{A})$ is a gadget as in our definition, using $U = W^{(M+1)}$, with accuracy ϵ (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

$$\eta = \omega_{M+1} = \frac{1}{2\|H\|}\left[\left(\frac{2\sqrt{2}}{\lambda_{\text{gap}} - \epsilon}\|H\| + 1\right)^M - 1\right]\left[\epsilon + (\zeta + 8\sqrt{2}\pi K\sqrt{\zeta})\|H\|\right] \ . \tag{3.152}$$

■

3.6.4 Gadget combination

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

Combination of general gadgets

Here we introduce some preparatory lemmas before proving Proposition 3.4.10.

The first lemma allows us to immediately reduce gadgets to the case that the unitary U is a direct rotation, defined in Definition 3.6.2. 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 3.6.7. *Suppose $(H', \mathcal{A})$ is a (η, ϵ)-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 [172] between the subspaces defined by $(\mathcal{I} \otimes P)$ and P'.

Proof of Lemma 3.6.7. By the gadget definition we have U such that

$$P' = U(\mathcal{I} \otimes P)U^\dagger \ , \quad \|U - \mathcal{I}\| \leq \eta \ . \tag{3.153}$$

As shown by Davis et al.[178], the direct rotation W between $(\mathcal{I} \otimes P)$ and P' minimises $\|W - \mathcal{I}\|$ subject to the first equality above, hence we have

$$P' = W(\mathcal{I} \otimes P)W^\dagger \ , \quad \|W - \mathcal{I}\| \leq \eta \ . \tag{3.154}$$

So

$$\begin{aligned} \|P'H'P' - W(H \otimes P)W^\dagger\| &\leq \epsilon + \|U(H \otimes P)U^\dagger - W(H \otimes P)W^\dagger\| \\ &\leq \epsilon + \|U(H \otimes P)U^\dagger - H \otimes P\| \\ &\quad + \|W(H \otimes P)W^\dagger - H \otimes P\| \\ &\leq \epsilon + 4\eta\|H\| \ . \end{aligned} \tag{3.155}$$

■

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 η small, P' will be close to the projector $\mathcal{I} \otimes P$. The following lemma provides a bound for H' when instead restricted to the span of $\mathcal{I} \otimes P$.

Lemma 3.6.8. *Suppose $(H', \mathcal{A})$ is a (η, ϵ)-gadget for H, with U, P', and P as in the definition, and where U is the direct rotation between $(\mathcal{I} \otimes P)$ and P'. Assume $\|H'\| \leq J'$, and*

$$\|(\mathcal{I} \otimes P)H'(\mathcal{I} \otimes P^\perp)\| \leq J'_O \ . \tag{3.156}$$

Then

$$\|(\mathcal{I} \otimes P)H'(\mathcal{I} \otimes P)\| \leq \|H\| + O(\epsilon + \eta J'_O + \eta^2 J') \ . \tag{3.157}$$

Proof of Lemma 3.6.8. Let S be the generator of the direct rotation U, so that $\|S\| = O(\eta)$ (see (3.100)) satisfies

$$P' = e^S(\mathcal{I}\otimes P)e^{-S} \ . \tag{3.158}$$

Hence

$$\|(\mathcal{I}\otimes P)H'(\mathcal{I}\otimes P)\| \leq \|(\mathcal{I}\otimes P)(e^{-S}H'e^S - H')(\mathcal{I}\otimes P)\| + \|P'H'P'\| \ . \tag{3.159}$$

By the gadget definition, we have $\|P'H'P'\| \leq \|H\| + \epsilon$, and using Lemma 3.6.5 we can bound the first term as

$$\|(\mathcal{I}\otimes P)(e^{-S}H'e^S - H')(\mathcal{I}\otimes P)\| \leq \|(\mathcal{I}\otimes P)[S,H'](\mathcal{I}\otimes P)\| + O(\eta^2 J') \ . \tag{3.160}$$

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

$$\|(\mathcal{I}\otimes P)[S,H'](\mathcal{I}\otimes P)\| \leq O(\eta J'_O) \ , \tag{3.161}$$

completing the proof. ∎

We now have the necessary tools to prove the gadget combination result Proposition 3.4.10. 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.[135], for which which similar techniques are used.

Proof of Proposition 3.4.10. We begin by reducing to the case of the direct rotation. By Lemma 3.6.7, we may replace ϵ with

$$\tilde{\epsilon} = \epsilon + 4\eta J = O(\epsilon + \eta J) \ , \tag{3.162}$$

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

$$\|(\mathcal{I}\otimes P_i)e^{-S_i}H'_i e^{S_i}(\mathcal{I}\otimes P_i) - H_i\otimes P_i\| \leq \tilde{\epsilon} \ , \tag{3.163}$$

where S_i is the generator of the direct rotation between the projectors $\mathcal{I}\otimes P_i$ and $P'_i := e^{S_i}(\mathcal{I}\otimes P_i)e^{-S_i}$. 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 $\mathcal{I}\otimes P_i$ and P'_i [172].

We define the operators

$$P := \otimes_i P_i \in \mathrm{Proj}(\otimes_i \mathcal{A}_i) \ , \quad S = \sum_i S_i \ . \tag{3.164}$$

We now use the triangle inequality along with Lemmas 3.6.5-3.6.6 to bound

$$\begin{aligned}
&\|(\mathcal{I}\otimes P)e^{-S}H'e^{S}(\mathcal{I}\otimes P) - H\otimes P\| \\
&\quad\leq \sum_i \|(\mathcal{I}\otimes P)(H_i' - [S,H_i'] + \frac{1}{2}[S,[S,H_i']] - \frac{1}{6}[S,[S,[S,H_i']]])(\mathcal{I}\otimes P) \\
&\quad - H_i\otimes P\| + O(n\eta^4 J') \ . \qquad (3.165)
\end{aligned}$$

Now we bound the terms in the norm separately, using that S_j is block-off-diagonal with respect to $(\mathcal{I}\otimes P_j)$, and that $(\mathcal{I}\otimes P_j)$ 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 $\mathcal{H}\otimes\mathcal{A}_k$.

- $(\mathcal{I}\otimes P)[S,H_i'](\mathcal{I}\otimes P)$:

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

  $$(\mathcal{I}\otimes P)[S,H_i'](\mathcal{I}\otimes P) = (\mathcal{I}\otimes P)[S_i,H_i'](\mathcal{I}\otimes P) \ . \qquad (3.166)$$

- $(\mathcal{I}\otimes P)[S,[S,H_i']](\mathcal{I}\otimes P)$:

  Note that if $j\neq k$, then

  $$(\mathcal{I}\otimes P)[S_j,[S_k,H_i']](\mathcal{I}\otimes P) = 0 \ , \qquad (3.167)$$

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

  $$(\mathcal{I}\otimes P)[S_j,[S_j,H_i']](\mathcal{I}\otimes P) \ . \qquad (3.168)$$

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

  $$\begin{aligned}
  \|(\mathcal{I}\otimes P)[S_j,[S_j,H_i']](\mathcal{I}\otimes P)\| &\leq 4\|S_j\|^2\|(\mathcal{I}\otimes P_i)H_i'(\mathcal{I}\otimes P_i)\| \\
  &\leq O(\eta^2 J + \eta^3 J_O' + \eta^4 J') \ , \qquad (3.169)
  \end{aligned}$$

  using Lemma 3.6.8, and neglecting the $O(\eta^2\epsilon)$ term which is dominated by the other terms in the regime of small ϵ to which the proposition applies. Hence

  $$(\mathcal{I}\otimes P)[S,[S,H_i']](\mathcal{I}\otimes P) = (\mathcal{I}\otimes P)[S_i,[S_i,H_i']](\mathcal{I}\otimes P) + O(\eta^2 J + \eta^3 J_O' + \eta^4 J') \ . \qquad (3.170)$$

  Here we have used the assumptions of bounded locality and degree in Setup 9. This ensures that only $O(1)$ of the local terms in S appear in the commutator $[S,H_i']$, and similarly for $[S,[S,H_i']]$.

- $(\mathcal{I}\otimes P)[S,[S,[S,H_i']]](\mathcal{I}\otimes P)$:

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

$$\begin{aligned}\|(\mathcal{I}\otimes P)[S_j,[S_k,[S_l,H_i']]](\mathcal{I}\otimes P)\| &\leq 8\|S_j\|\|S_k\|\|S_l\|\|(\mathcal{I}\otimes P_i)H_i'(\mathcal{I}\otimes P_i)\| \\ &\leq O(\eta^3 J+\eta^4 J_O'+\eta^5 J')\ , \end{aligned} \tag{3.171}$$

by Lemma 3.6.8, 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 $(\mathcal{I}\otimes P_j)$ 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 $(\mathcal{I}\otimes P_i)S_i = S_i(\mathcal{I}\otimes P_i^{\perp})$, we arrive at

$$\begin{aligned}\|(\mathcal{I}\otimes P)[S_j,[S_k,[S_l,H_i']]](\mathcal{I}\otimes P)\| &\leq O(\eta^3)\|(\mathcal{I}\otimes P_i^{\perp})H_i'(\mathcal{I}\otimes P_i)\| \\ &= O(\eta^3 J_O')\ . \end{aligned} \tag{3.172}$$

Hence

$$\begin{aligned}(\mathcal{I}\otimes P)[S,[S,[S,H_i']]](\mathcal{I}\otimes P) &= (\mathcal{I}\otimes P)[S_i,[S_i,[S_i,H_i']]](\mathcal{I}\otimes P) \\ &\quad + O(\eta^3 J+\eta^3 J_O'+\eta^5 J')\ , \end{aligned} \tag{3.173}$$

once again using the locality assumptions of Setup 9.

So putting the above bounds together and applying Lemma 3.6.5, we obtain

$$\begin{aligned}\|(\mathcal{I}\otimes P)e^{-S}H'e^{S}(\mathcal{I}\otimes P)-H\otimes P\| &\leq \sum_i \|(\mathcal{I}\otimes P)e^{-S_i}H_i'e^{S_i}(\mathcal{I}\otimes P)-H_i\otimes P\| \\ &\quad + O(n\eta^2 J+n\eta^3 J_O'+n\eta^4 J') \\ &\leq O(n\epsilon+n\eta J+n\eta^3 J_O'+n\eta^4 J')\ , \end{aligned} \tag{3.174}$$

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

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

■

Combination of low-energy gadgets

Having proved Proposition 3.4.10, we know that the H' from Proposition 3.4.11 is an (η',ϵ')-gadget for H. It remains to prove that it is in fact a $(\Delta',\eta',\epsilon')$-gadget, which requires replacing the projector $\tilde{P}=e^S(\mathcal{I}\otimes P)e^{-S}$ in (3.174) 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 3.6.4).

Lemma 3.6.9. *Let $A \in \mathrm{Herm}(\mathcal{H})$, and let $P \in \mathrm{Proj}(\mathcal{H})$ be a projector of the same rank as $P_{\leq\Delta(A)}$, where $P_{\leq\Delta(A)} \in \mathrm{Proj}(\mathcal{H})$ is the projector onto the eigenvectors of A with eigenvalues less than Δ. Suppose that $\|PAP\| \leq \lambda$.*

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

$$\|U - \mathcal{I}\| \leq \frac{\sqrt{2}}{\Delta - \lambda}\|P^\perp AP\| \ . \tag{3.175}$$

Proof of Lemma 3.6.9. Follows from Lemma 3.6.4 using $A \mapsto PAP$, $B \mapsto A$. ∎

Proof of Proposition 3.4.11. By Proposition 3.4.10, we have

$$\tau := \|(\mathcal{I}\otimes P)e^{-S}H'e^{S}(\mathcal{I}\otimes P) - H\otimes P\| \leq O(n\epsilon + n\eta J + n\eta^3 J'_O + n\eta^4 J') \ , \tag{3.176}$$

and by condition (3.48) this implies that

$$\frac{\tau}{\Delta} \to 0 \quad \text{as } n \to \infty. \tag{3.177}$$

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

$$\|\tilde{P}H'\tilde{P}\| \leq \|H\| + \tau \ . \tag{3.178}$$

We seek to use Lemma 3.6.9 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

$$\begin{aligned}\|\tilde{P}H'\tilde{P}^\perp\| &= \|(\mathcal{I}\otimes P)e^{-S}H'e^{S}(\mathcal{I}\otimes P^\perp)\| \\ &\leq \sum_i \|(\mathcal{I}\otimes P)e^{-S}H'_i e^{S}(\mathcal{I}\otimes P^\perp)\| \ ,\end{aligned} \tag{3.179}$$

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 $(\mathcal{I}\otimes P_i)$, in which S_i is block off-diagonal. Here we are using the definition of the direct rotation, and the fact that $e^{S_i}(\mathcal{I}\otimes P_i)e^{-S_i}$ 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

$$\|\tilde{P}H'\tilde{P}^\perp\| \leq \sum_i \|(\mathcal{I}\otimes P)e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^{S}(\mathcal{I}\otimes P^\perp)\| \ . \tag{3.180}$$

From here we use Lemma 3.6.5 to expand

$$e^{S_i}\tilde{H}_i e^{-S_i} = \tilde{H}_i + [S_i, \tilde{H}_i] + R_i \ , \tag{3.181}$$

where R_i acts on $O(1)$ sites, and by Lemma 3.6.6 we can bound $\|R_i\| = O(J'\eta^2)$. Hence

$$\begin{aligned}e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^{S} = \tilde{H}_i &+ [S_i, \tilde{H}_i] + R_i \\ &- [S, \tilde{H}_i + [S_i, \tilde{H}_i] + R_i] \\ &+ \tilde{R}_i \ ,\end{aligned} \tag{3.182}$$

where the remainder $\tilde{R}_i$ is similarly obtained by Lemmas 3.6.5-3.6.6 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

$$e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^{S} = \tilde{H}_i - \sum_{j\neq i}[S_j, \tilde{H}_i] + O(J'\eta^2) \ . \tag{3.183}$$

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

$$\begin{aligned}\|(\mathcal{I}\otimes P)[S_j, \tilde{H}_i](\mathcal{I}\otimes P^{\perp})\| &\leq 2\|S_j\|\|(\mathcal{I}\otimes P_i)\tilde{H}_i\| \\ &\leq O(\eta)\|e^{-S_i}(\mathcal{I}\otimes P_i)e^{S_i}H_i'\| \\ &= O(\eta)\|P_{\leq\Delta(H_i')}H_i'\| \\ &\leq O(\eta)(\|H_i\| + \epsilon) = O(\eta J + \eta\epsilon) \ . \end{aligned} \tag{3.184}$$

In the last line we have used the fact that H' defines a (Δ, η, ϵ)-gadget for H_i. Hence we can conclude that

$$(\mathcal{I}\otimes P)e^{-S}e^{S_i}\tilde{H}_i e^{-S_i}e^{S}(\mathcal{I}\otimes P^{\perp}) = O(J'\eta^2 + \eta J + \eta\epsilon) \ , \tag{3.185}$$

and so, inserting into (3.179), we have

$$\|\tilde{P}H'\tilde{P}^{\perp}\| := \omega = O(nJ'\eta^2 + n\eta J + n\eta\epsilon) \ . \tag{3.186}$$

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

  $$\begin{aligned}&(\langle\psi_{\mathcal{H}}|\otimes\langle\psi_{\mathcal{A}_i}|)H_i'(|\psi_{\mathcal{H}}\rangle\otimes|\psi_{\mathcal{A}_i}\rangle) \\ &\quad\geq (\langle\psi|_{\mathcal{H}}\otimes\langle\psi_{\mathcal{A}_i}|)P_{\leq\Delta(H_i')}H_i'P_{\leq\Delta(H_i')}(|\psi_{\mathcal{H}}\rangle\otimes|\psi_{\mathcal{A}_i}\rangle) \\ &\quad\geq (\langle\psi|_{\mathcal{H}}\otimes\langle\psi_{\mathcal{A}_i}|)e^{S_i}(H_i\otimes P_i)e^{-S_i}(|\psi_{\mathcal{H}}\rangle\otimes|\psi_{\mathcal{A}_i}\rangle) - \epsilon \\ &\quad\geq \langle\psi_{\mathcal{H}}|H_i|\psi_{\mathcal{H}}\rangle - (\epsilon + 2J\eta) \ . \end{aligned} \tag{3.187}$$

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

  Then

  $$\begin{aligned}&(\langle\psi_{\mathcal{H}}|\otimes\langle\psi_{\mathcal{A}_i}|)H_i'(|\psi_{\mathcal{H}}\rangle\otimes|\psi_{\mathcal{A}_i}\rangle) \\ &\quad\geq \Delta(\langle\psi_{\mathcal{H}}|\otimes\langle\psi_{\mathcal{A}_i}|)P_{>\Delta(H_i')}(|\psi_{\mathcal{H}}\rangle\otimes|\psi_{\mathcal{A}_i}\rangle) \\ &\quad= \Delta\big((\langle\psi_{\mathcal{H}}|\otimes\langle\psi_{\mathcal{A}_i}|)e^{S_i}(\mathcal{I}\otimes P_i^{\perp})e^{-S_i}(|\psi_{\mathcal{H}}\rangle\otimes|\psi_{\mathcal{A}_i}\rangle)\big) \\ &\quad\geq \Delta(1-2\eta) \\ &\quad\geq \langle\psi_{\mathcal{H}}|H_i|\psi_{\mathcal{H}}\rangle \ , \end{aligned} \tag{3.188}$$

  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

$$|\psi\rangle = e^S |\psi_{\mathcal{H}}\rangle \otimes (\otimes_i |\psi_{\mathcal{A}_i}\rangle) \ , \tag{3.189}$$

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

$$\begin{aligned}
\langle\psi|\tilde{P}^\perp H' \tilde{P}^\perp|\psi\rangle &= \sum_i \langle\psi|H_i'|\psi\rangle \\
&\geq \sum_{i\neq j} \big(\langle\psi_{\mathcal{H}}|H_i|\psi_{\mathcal{H}}\rangle - \epsilon - 2J\eta\big) + \Delta(1-2\eta) \\
&\geq \langle\psi_{\mathcal{H}}|H|\psi_{\mathcal{H}}\rangle - (N-1)(\epsilon+2J\eta) - J + \Delta(1-2\eta) \\
&\geq \Delta(1-2\eta) - \big(\|H\| + J + N(\epsilon+2J\eta)\big) \\
&\geq \frac{3}{4}\Delta \ ,
\end{aligned} \tag{3.190}$$

where we have used the condition (3.46).

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 (3.186). Based on (3.178) and (3.190), we know that

$$\operatorname{spec} \tilde{H} \subset [-(\|H\|+\tau), \|H\|+\tau] \cup [\frac{3}{4}\Delta, \infty) \ , \tag{3.191}$$

corresponding to low- and high-energy projectors $\tilde{P}$ and $\tilde{P}^\perp$ respectively. Note that (3.46) in particular implies that $\Delta \geq 4\|H\|$, so $\frac{1}{2}\Delta - \|H\| \geq \frac{1}{4}\Delta$. Using this, and (3.177), we see

$$\frac{\frac{1}{2}\Delta - (\|H\|+\tau)}{\Delta} \geq \frac{1}{4} - \frac{\tau}{\Delta} = \Omega(1) \tag{3.192}$$

for large n. But by (3.46), (3.47), and (3.48),

$$\frac{\omega}{\Delta} = O(n\eta^2 J'/\Delta) + O(n\eta J/\Delta) + O(n\eta\epsilon/\Delta) = o(1) \ . \tag{3.193}$$

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

$$\operatorname{spec} \tilde{H} \subset (-\infty, \frac{1}{2}\Delta - \omega] \cup [\frac{3}{4}\Delta, \infty) \ , \tag{3.194}$$

once again corresponding to subspaces defined by $\tilde{P}$ and $\tilde{P}^\perp$. Now, by (3.186) and Lemma 3.6.1, 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

$$\operatorname{spec} H' \subset (-\infty, \frac{1}{2}\Delta] \cup [\frac{3}{4}\Delta - \omega, \infty) \ . \tag{3.195}$$

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 Δ' in the image of $\tilde{P}^\perp$, which is disallowed by (3.191). So the direct rotation $W \in \mathrm{U}(\mathcal{H}\otimes\mathcal{A})$ from $\tilde{P}$ to $P_{\leq\Delta'(H')}$ is well-defined, and by Lemma 3.6.9

$$\|W - \mathcal{I}\| \leq \frac{\sqrt{2}}{\frac{1}{2}\Delta - \|H\| - \tau}\omega \ . \tag{3.196}$$

Again using that $\Delta \geq 4\|H\|$, so $\frac{1}{2}\Delta - \|H\| \geq \frac{1}{4}\Delta = \Omega(\Delta)$. By (3.48), this will dominate the relatively small τ term, so using (3.186) and that $\Delta = \Omega(J')$ we have

$$\begin{aligned}\|W - \mathcal{I}\| &= O(\omega/J') = O(n\eta^2 + n\eta J/J' + n\eta\epsilon/J') \\ &= O\big(n\eta^2(1 + J/\eta J' + \epsilon/J'\eta)\big) \\ &= O\Big(n\eta^2\big(1 + o(1/n) + o(\epsilon/nJ)\big)\Big) = O(n\eta^2) \ , \end{aligned} \tag{3.197}$$

where in the last line we used (3.48). 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 (3.100). Then

$$P_{\leq \Delta'(H')} = U(\mathcal{I} \otimes P)U^\dagger \ , \tag{3.198}$$

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

$$\begin{aligned}\|U - \mathcal{I}\| &\leq \|e^X(e^S - \mathcal{I})\| + \|e^X - \mathcal{I}\| \\ &\leq \|S\| + \|X\| \\ &= O(n\eta) + O(n\eta^2) = O(n\eta) \ . \end{aligned} \tag{3.199}$$

It remains to bound find ϵ' to achieve a bound of the form

$$\|P_{\leq \Delta'(H')} H' P_{\leq \Delta'(H')} - U(H \otimes P)U^\dagger\| \leq \epsilon' \ . \tag{3.200}$$

Using (3.178) and the triangle inequality we have

$$\|P_{\leq \Delta'(H')} H' P_{\leq \Delta'(H')} - U(H \otimes P)U^\dagger\| \leq \|\tilde{P}(e^{-X} H' e^X - H')\tilde{P}\| + \tau \ , \tag{3.201}$$

and the first term can be bounded using Lemma 3.6.5 (and bounding the remainder with (3.118)) by

$$\begin{aligned}\|\tilde{P}(e^{-X} H' e^X - H')\tilde{P}\| &\leq \|\tilde{P}[X, H']\tilde{P}\| + \frac{1}{2}\|[X, [X, H']]\| \\ &\leq 2\|X\| \cdot \|\tilde{P}H'\tilde{P}^\perp\| + 2\|X\|^2 \cdot \|H'\| \\ &\leq O(n^3\eta^4 J' + n^2\eta^3 J) \ . \end{aligned} \tag{3.202}$$

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 (3.186) to bound $\|\tilde{P}H'\tilde{P}^\perp\|$. Hence

$$\epsilon' = O(n\epsilon + n\eta J + n\eta^3 J'_O + n^3\eta^4 J') \ . \tag{3.203}$$

■

Ground-state estimation with bounded-strength low-energy gadgets

The following proof of Theorem 3.4.12 is a simple corollary of Proposition 3.4.10, and generalises the proof of Bravyi et al., Theorem 1[135].

Proof of Theorem 3.4.12. For the first part of this proof, we seek to put a lower bound on the individual gadgets H_i'. We write

$$H_i' = P_{\leq\Delta(H_i')} H_i' P_{\leq\Delta(H_i')} + P_{>\Delta(H_i')} H_i' P_{>\Delta(H_i')} \ , \tag{3.204}$$

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

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

$$\begin{aligned} \|P_{\leq\Delta(H_i')} H_i' P_{\leq\Delta(H_i')} - H_i \otimes P_i\| &\leq \epsilon + \|e^{S_i}(H_i \otimes P_i)e^{-S_i} - H_i \otimes P_i\| \\ &\leq O(\epsilon + \eta J) \ , \end{aligned} \tag{3.205}$$

using Lemma 3.6.5, and hence

$$P_{\leq\Delta(H_i')} H_i' P_{\leq\Delta(H_i')} \geq H_i \otimes P_i + O(\epsilon + \eta J) \ . \tag{3.206}$$

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

$$P_{>\Delta(H_i')} H_i' P_{>\Delta(H_i')} \geq P_{>\Delta(H_i')}(H_i \otimes \mathcal{I}) P_{>\Delta(H_i')} + O(\epsilon) \ . \tag{3.207}$$

Furthermore, we can approximate the RHS of this expression by

$$\begin{aligned} &\|P_{>\Delta(H_i')}(H_i \otimes \mathcal{I}) P_{>\Delta(H_i')} - H_i \otimes P_i^{\perp}\| \\ &\quad = \|e^{S_i}(\mathcal{I} \otimes P_i^{\perp})e^{-S_i}(H_i \otimes \mathcal{I})e^{S_i}(\mathcal{I} \otimes P_i^{\perp})e^{-S_i} - H_i \otimes P_i^{\perp}\| \\ &\quad \leq O(\eta J) \ , \end{aligned} \tag{3.208}$$

by applying Lemma 3.6.5. Hence

$$P_{>\Delta(H_i')} H_i' P_{>\Delta(H_i')} \geq H_i \otimes P_i^{\perp} + O(\epsilon + \eta J) \ . \tag{3.209}$$

Summing (3.206) and (3.209) for all i, we obtain

$$H' = \sum_i H_i' \geq \sum_i (H_i \otimes \mathcal{I} + O(\epsilon + \eta J)) = H \otimes \mathcal{I} + O(n\epsilon + n\eta J) \ , \tag{3.210}$$

and so the ground state energy of H' must satisfy

$$\lambda_0(H') \geq \lambda_0(H) + O(n\epsilon + n\eta J) \ . \tag{3.211}$$

Now notice that the restriction of H' to a subspace can only increase λ_0, so

$$\lambda_0(H') = \lambda_0(e^{-S} H' e^{S}) \leq \lambda_0((\mathcal{I} \otimes P)e^{-S} H' e^{S}(\mathcal{I} \otimes P)) \ . \tag{3.212}$$

Here, as in the work of Bravyi et al.[135], we abuse notation slightly: in the expression $\lambda_0((\mathcal{I} \otimes P)e^{-S} H' e^{S}(\mathcal{I} \otimes P))$, we implicitly take the ground state of the restriction of $e^{-S} H' e^{S}$ to the image of $(\mathcal{I} \otimes P)$. So by Proposition 3.4.10 we have

$$\lambda_0(H') \leq \lambda_0(H) + O(n\epsilon + n\eta J + n\eta^3 J_O' + n\eta^4 J') \ . \tag{3.213}$$

Combining (3.211)-(3.213) gives the desired result. ∎

3.6.5 Gadget energy scaling

In this section, we prove Theorem 3.4.13. 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 3.6.10 (k-local function)**.** Let $f : \{0,1\}^n \to \mathbb{R}$ be a function. We say that f is *k-local* if it can be written as a sum of functions

$$f(x_1, x_2, \ldots, x_n) = \sum_i f_i(x_1, x_2, \ldots, x_n) \ , \tag{3.214}$$

where the $f_i : \{0,1\}^n \to \mathbb{R}$ 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 3.6.11. *Let f be a k-local function on n inputs. Then $\mathcal{R}f : \{0,1\}^{n-1} \to \mathbb{R}$, defined by*

$$\mathcal{R}f(x_1, \ldots, x_{n-1}) = f(x_1, \ldots, x_{n-1}, 0) - f(x_1, \ldots, x_{n-1}, 1) \tag{3.215}$$

is $(k-1)$-local.

Proof of Lemma 3.6.11. Decomposing $f = \sum_i f_i$ as in Definition 3.6.10, 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 3.6.12. *Let $k > k' > 0$. There exists a k-local function $f : \{0,1\}^k \to \mathbb{R}$ with $\max_{x \in \{0,1\}^k} |f(x)| \leq 1$ such that for any k'-local function $g : \{0,1\}^k \to \mathbb{R}$,*

$$\max_{x \in \{0,1\}^k} |f(x) - g(x)| \geq 1 \ . \tag{3.216}$$

Proof of Lemma 3.6.12. For any $r \geq 1$, we can define $\mathcal{R}^r f : \{0,1\}^{k-r} \to \mathbb{R}$ by

$$\mathcal{R}^r f(x_1, \ldots, x_{k-r}) = \sum_{x_{k-r+1}, \ldots, x_k \in \{0,1\}} (-1)^{\sum_{j=1}^r x_{k-r+j}} f(x_1, \ldots, x_k) \ . \tag{3.217}$$

Applying Lemma 3.6.11 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 $f : \{x_1, \ldots, x_k\}$ be the parity function

$$f(x_1, \ldots, x_k) = (-1)^{\sum_{i=1}^k x_i} \ . \tag{3.218}$$

Then we can calculate

$$\begin{aligned} \mathcal{R}^{k'} f(x_1, \ldots, x_{k-k'}) &= \sum_{x_{k-k'+1}, \ldots, x_k \in \{0,1\}} (-1)^{\sum_{j=1}^{k'} x_{k-k'+j}} \cdot (-1)^{\sum_{i=1}^k x_i} \\ &= 2^{k'} (-1)^{\sum_{j=1}^{k-k'} x_j} \end{aligned} \tag{3.219}$$

Hence $\mathcal{R}^{k'} f(x_1, \ldots, x_{k-k'})$ 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

$$|\mathcal{R}^{k'} f(y) - \mathcal{R}^{k'} g(y)| \geq 2^{k'} \ . \tag{3.220}$$

Hence, expanding $\mathcal{R}^{k'} f(y)$ and $\mathcal{R}^{k'} g(y)$ into $2^{k'}$ terms using (3.217), we must have

$$\max_{x \in \{0,1\}^k} |f(x) - g(x)| \geq 1 \ . \tag{3.221}$$

■

Note that Lemma 3.6.12 uses the parity function (which appears in the proof of Theorem 3.4.13) 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 3.6.12 to argue that the target k-local term cannot be reproduced by a k'-local Hamiltonian.

Proof of Theorem 3.4.13. By the gadget definition, we have $U \in \mathrm{U}(\mathcal{H} \otimes \mathcal{A})$ and $P \in \mathrm{Proj}(\mathcal{A})$ such that

$$\|U - \mathcal{I}\| \leq \eta \ , \quad \|P' H' P' - U(H \otimes P)U^\dagger\| \leq \epsilon \ , \quad \text{where } P' = U(\mathcal{I} \otimes P)U^\dagger \ . \tag{3.222}$$

For any given $x \in \{0,1\}^k$, define $|\psi_x\rangle \in \mathcal{H}$ to be the pure state whose ith 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 $F, f : \{0,1\}^k \to \mathbb{R}$ by

$$\begin{aligned} F(x) &= \mathrm{tr}[H|\psi_x\rangle\langle\psi_x|] \ , \\ f(x) &= \mathrm{tr}[H'\big(|\psi_x\rangle\langle\psi_x| \otimes |\phi\rangle\langle\phi|\big)] \ . \end{aligned} \tag{3.223}$$

Notice that F and f are k- and k'-local respectively, and by Lemma 3.6.12 there exists some $y \in \{0,1\}^k$ such that

$$|F(y) - f(y)| \geq J \ . \tag{3.224}$$

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

$$\begin{aligned} |F(x) - f(x)| &= |\,\mathrm{tr}[(H \otimes P)(|\psi_x\rangle\langle\psi_x| \otimes |\phi\rangle\langle\phi|)] - \mathrm{tr}[H'(|\psi_x\rangle\langle\psi_x| \otimes |\phi\rangle\langle\phi|)]| \\ &\leq \|(\mathcal{I} \otimes P)H'(\mathcal{I} \otimes P) - H \otimes P\| \\ &= \|P'UH'U^\dagger P' - U(H \otimes P)U^\dagger\| \\ &\leq \epsilon + \|P'(H' - UH'U^\dagger)P'\| \\ &\leq \epsilon + 2\eta\|H'\| \ . \end{aligned} \tag{3.225}$$

Hence we must have scaling

$$\|H'\| \geq \frac{J - \epsilon}{2\eta} \ , \tag{3.226}$$

as required. ■

3.6.6 Dissipative gadgets

Dissipative gadgets in isolation

Here we first prove Proposition 3.4.14. This follows from direct calculation, by Taylor expanding the expression $e^{-i\delta t H'}$ and identifying the leading order terms in δt. 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 3.4.14. First, notice that the requirement $H^2_{|1\rangle\langle 1|} = \omega^2 \mathcal{I}$ implies that

$$e^{-i\delta t H_{|1\rangle\langle 1|}} = \mathcal{I} \ , \quad H^{-1}_{|1\rangle\langle 1|} = \omega^{-2} H_{|1\rangle\langle 1|} \ . \tag{3.227}$$

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

$$e^{-i\delta t H'} = \sum_{k\geq 0} \frac{(-i\delta t)^k}{k!} \big(H_{\mathcal{I}} \otimes \mathcal{I} + H_X \otimes X + H_{|1\rangle\langle 1|} \otimes |1\rangle\langle 1|\big)^k. \tag{3.228}$$

We can expand out this expression so that each term is a product of a factors of $H_{\mathcal{I}} \otimes \mathcal{I}$, 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. 3.11.

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 3.11: 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 (3.228). We denote $\mathbb{N}_+ = \{1, 2, \dots\}$ and $\mathbb{N} = \{0, 1, 2, \dots\}$.

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

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

  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 $(\mathcal{I}\otimes\langle 0|)e^{-i\delta tH'}(\mathcal{I}\otimes|0\rangle)$ from Fig. 3.11 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

$$(\mathcal{I}\otimes\langle 0|)e^{-i\delta tH'}(\mathcal{I}\otimes|0\rangle) = \mathcal{I} - i\delta t H_{\mathcal{I}} + \sum_{k\geq 2}\frac{(-i\delta t)^k}{k!}H_X H_{|1\rangle\langle 1|}^{k-2}H_X + O((\delta t)^2)\ . \tag{3.229}$$

Furthermore, the sum can be simplified to

$$\begin{aligned}\sum_{k\geq 2}\frac{(-i\delta t)^k}{k!}H_X H_{|1\rangle\langle 1|}^{k-2}H_X &= H_X H_{|1\rangle\langle 1|}^{-2}\Big(\sum_{k\geq 2}\frac{(-i\delta t)^k}{k!}H_{|1\rangle\langle 1|}^{k}\Big)H_X \\ &= H_X H_{|1\rangle\langle 1|}^{-2}\Big(e^{-i\delta tH_{|1\rangle\langle 1|}} - \mathcal{I} + i\delta t H_{|1\rangle\langle 1|}\Big)H_X \\ &= i\delta t\omega^{-2}H_X H_{|1\rangle\langle 1|}H_X\ ,\end{aligned} \tag{3.230}$$

using (3.227). Hence we have shown that

$$(\mathcal{I}\otimes\langle 0|)e^{-i\delta tH'}(\mathcal{I}\otimes|0\rangle) = \mathcal{I} - i\delta t(H_{\mathcal{I}} - \omega^{-2}H_X H_{|1\rangle\langle 1|}H_X) + O((\delta t)^2)\ . \tag{3.231}$$

- $(\mathcal{I}\otimes\langle 1|)e^{-i\delta tH'}(\mathcal{I}\otimes|0\rangle)$:

  By a similar argument to above, the only contributing terms from Fig. 3.11 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

$$(\mathcal{I}\otimes\langle 1|)e^{-i\delta tH'}(\mathcal{I}\otimes|0\rangle) = \sum_{k\geq 1}\frac{(-i\delta t)^k}{k!}H_{|1\rangle\langle 1|}^{k-1}H_X + O((\delta t)^{3/2})\ . \tag{3.232}$$

  The sum here can be similarly simplified by (3.227):

$$\begin{aligned}\sum_{k\geq 1}\frac{(-i\delta t)^k}{k!}H_{|1\rangle\langle 1|}^{k-1}H_X &= H_{|1\rangle\langle 1|}^{-1}\Big(\sum_{k\geq 0}\frac{(-i\delta t)^k}{k!}H_{|1\rangle\langle 1|}^{k} - \mathcal{I}\Big)H_X \\ &= H_{|1\rangle\langle 1|}^{-1}\big(e^{-i\delta tH_{|1\rangle\langle 1|}} - \mathcal{I}\big)H_X \\ &= 0\ ,\end{aligned} \tag{3.233}$$

  so

$$(\mathcal{I}\otimes\langle 1|)e^{-i\delta tH'}(\mathcal{I}\otimes|0\rangle) = O((\delta t)^{3/2})\ , \tag{3.234}$$

  which, along with (3.231), completes the proof.

■

Trotter errors

The idea for the proof of Proposition 3.4.15 is to factorise the overall evolution operator

$$e^{-i\delta t(H'+H_{\text{else}}\otimes\mathcal{I})} \approx e^{-i\delta t H_{\text{else}}\otimes\mathcal{I}} e^{-i\delta t H'} \ . \tag{3.235}$$

From here, we simply apply Proposition 3.4.14 to the initial state $|\psi\rangle \otimes |0\rangle$, and then evolve the $\mathcal{H}$ system of the resultant state under H_{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 δt, under which only a limited set of interactions in H_{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.[179]. 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

$$\begin{aligned} f'(t) &= Ae^{tA}e^{tB} + e^{tA}Be^{tB} \\ &= (A+B)f(t) + e^{tA}\big(B - e^{-tA}Be^{tA}\big)e^{tB} \ . \end{aligned} \tag{3.236}$$

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

Lemma 3.6.13 (Variation of parameters formula [179])**.** *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*

$$f'(t) = Kf(t) + L(t) \ , \quad f(0) = \mathcal{I} \ . \tag{3.237}$$

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

$$f(t) = e^{tK} + \int_0^t \mathrm{d}\tau e^{(t-\tau)K} L(\tau) \ . \tag{3.238}$$

Hence, using Lemma 3.6.13 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

$$e^{tA}e^{tB} - e^{t(A+B)} = \int_0^t \mathrm{d}\tau e^{(t-\tau)(A+B)} e^{\tau A}\big(B - e^{-\tau A}Be^{\tau A}\big)e^{\tau B} \tag{3.239}$$

The expression (3.239) 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 3.6.14. *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*

$$\|e^{itH} A e^{-itH} - A\| \leq O(\|A\| t) \ . \tag{3.240}$$

Proof of Lemma 3.6.14. For $X \in \mathrm{Lin}(\mathcal{H})$, define $f_X(t) = \mathrm{tr}[X(e^{itH} A e^{-itH} - A)]$, so that

$$\|e^{itH} A e^{-itH} - A\| = \max_{X \in D(\mathcal{H})} |f_X(t)| \ . \tag{3.241}$$

We can see that $f_X(0) = 0$ and $f_X'(t) = i\,\mathrm{tr}[X e^{itH}[H, A] e^{-itH}]$. 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

$$f_X(t) = O(\|X\|_1 \|A\| t) \ , \tag{3.242}$$

so by (3.241) we are done. ∎

Dissipative gadgets with other terms

Proof of Proposition 3.4.15. Using (3.239) with $t = \delta t$, $A = -iH_{\mathrm{else}} \otimes \mathcal{I}$, and $B = -iH'$, we obtain a Trotter error given by

$$\begin{aligned} &e^{-i\delta t H_{\mathrm{else}} \otimes \mathcal{I}} e^{-i\delta t H'} - e^{-i\delta t (H_{\mathrm{else}} \otimes \mathcal{I} + H')} := E \\ &\quad = -i \int_0^{\delta t} \mathrm{d}\tau e^{-i(\delta t - \tau)(H_{\mathrm{else}} \otimes \mathcal{I} + H')} e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}} \big(H' - e^{i\tau H_{\mathrm{else}} \otimes \mathcal{I}} H' e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}}\big) e^{-i\tau H'} \ . \end{aligned} \tag{3.243}$$

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

$$E = \begin{pmatrix} (\mathcal{I} \otimes \langle 0|) E (\mathcal{I} \otimes |0\rangle) & (\mathcal{I} \otimes \langle 0|) E (\mathcal{I} \otimes |1\rangle) \\ (\mathcal{I} \otimes \langle 1|) E (\mathcal{I} \otimes |0\rangle) & (\mathcal{I} \otimes \langle 1|) E (\mathcal{I} \otimes |1\rangle) \end{pmatrix} \ , \tag{3.244}$$

and we will focus on individually bounding these blocks. Notice that, commuting projectors on the ancillary space past $H_{\mathrm{else}} \otimes \mathcal{I}$ and applying Lemma 3.6.14, we have

$$\begin{aligned} &\|(\mathcal{I} \otimes \langle 0|)(H' - e^{i\tau H_{\mathrm{else}} \otimes \mathcal{I}} H' e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}})(\mathcal{I} \otimes |0\rangle)\| \\ &\quad = \|(\mathcal{I} \otimes \langle 0|) H' (\mathcal{I} \otimes |0\rangle) - e^{i\tau H_{\mathrm{else}}} (\mathcal{I} \otimes \langle 0|) H' (\mathcal{I} \otimes |0\rangle) e^{-i\tau H_{\mathrm{else}}}\| \\ &\quad = \|H_{\mathcal{I}} - e^{i\tau H_{\mathrm{else}}} H_{\mathcal{I}} e^{-i\tau H_{\mathrm{else}}}\| = O(\delta t) \ . \end{aligned} \tag{3.245}$$

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

$$\|(\mathcal{I} \otimes \langle 0|)(H' - e^{i\tau H_{\mathrm{else}} \otimes \mathcal{I}} H' e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}})(\mathcal{I} \otimes |0\rangle)\| = O(\delta t) \ , \tag{3.246a}$$

$$\|(\mathcal{I} \otimes \langle 1|)(H' - e^{i\tau H_{\mathrm{else}} \otimes \mathcal{I}} H' e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}})(\mathcal{I} \otimes |0\rangle)\| = O(\delta t^{1/2}) \ , \tag{3.246b}$$

$$\|(\mathcal{I} \otimes \langle 0|)(H' - e^{i\tau H_{\mathrm{else}} \otimes \mathcal{I}} H' e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}})(\mathcal{I} \otimes |1\rangle)\| = O(\delta t^{1/2}) \ , \tag{3.246c}$$

$$\|(\mathcal{I} \otimes \langle 1|)(H' - e^{i\tau H_{\mathrm{else}} \otimes \mathcal{I}} H' e^{-i\tau H_{\mathrm{else}} \otimes \mathcal{I}})(\mathcal{I} \otimes |1\rangle)\| = O(1) \ . \tag{3.246d}$$

Since our initial state is of the form $|\psi\rangle \otimes |0\rangle$, we need only bound the magnitudes of the blocks $(\mathcal{I}\otimes\langle 0|)E(\mathcal{I}\otimes|0\rangle)$ and $(\mathcal{I}\otimes\langle 1|)E(\mathcal{I}\otimes|0\rangle)$. To this end, we need to describe the action of the operator $e^{-i\tau H'}$ on the operators $(\mathcal{I}\otimes|0\rangle)$ and $(\mathcal{I}\otimes|1\rangle)$, for $0 \leq \tau \leq \delta t$. By considering the series expansion as in (3.228), 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

$$e^{-i\tau H'}(\mathcal{I}\otimes|0\rangle) = O(1)\otimes|0\rangle + O((\delta t)^{1/2})\otimes|1\rangle \ , \tag{3.247a}$$

$$e^{-i\tau H'}(\mathcal{I}\otimes|1\rangle) = O((\delta t)^{1/2})\otimes|0\rangle + O(1)\otimes|1\rangle \ . \tag{3.247b}$$

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:

$$\begin{aligned}\|E\| &\leq \int_0^{\delta t} \mathrm{d}\tau \|H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}} H' e^{-i\tau H_{\text{else}}\otimes\mathcal{I}}\| \\ &\leq \int_0^{\delta t} \mathrm{d}\tau O(1) = O(\delta t) \ ,\end{aligned} \tag{3.248}$$

using (3.246d) to give the most pessimistic bound. Therefore in particular

$$e^{i\tau H_{\text{else}}\otimes\mathcal{I}} e^{i(\delta t-\tau)(H_{\text{else}}\otimes\mathcal{I}+H')} = e^{i\delta t H_{\text{else}}\otimes\mathcal{I}} e^{i(\delta t-\tau)H'} + O(\delta t) \ , \tag{3.249}$$

so, using (3.247a),

$$\begin{aligned}e^{i\tau H_{\text{else}}\otimes\mathcal{I}} e^{i(\delta t-\tau)(H_{\text{else}}\otimes\mathcal{I}+H')}(\mathcal{I}\otimes|0\rangle) &= e^{-i\delta t H_{\text{else}}\otimes\mathcal{I}} e^{i(\delta t-\tau)H'}(\mathcal{I}\otimes|0\rangle) + O(\delta t) \\ &= O(1)\otimes|0\rangle + O((\delta t)^{1/2})\otimes|1\rangle \ ,\end{aligned} \tag{3.250}$$

and similarly

$$e^{i\tau H_{\text{else}}\otimes\mathcal{I}} e^{i(\delta t-\tau)(H_{\text{else}}\otimes\mathcal{I}+H')}(\mathcal{I}\otimes|1\rangle) = O((\delta t)^{1/2})\otimes|0\rangle + O(1)\otimes|1\rangle \ . \tag{3.251}$$

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

$$\begin{aligned}&\|(\mathcal{I}\otimes\langle 0|)E(\mathcal{I}\otimes|0\rangle)\| \\ &\quad\leq \int_0^{\delta t} \mathrm{d}\tau \|(\mathcal{I}\otimes\langle 0|)e^{-i(\delta t-\tau)(H_{\text{else}}\otimes\mathcal{I}+H')} e^{-i\tau H_{\text{else}}\otimes\mathcal{I}}(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}} H' e^{-i\tau H_{\text{else}}\otimes\mathcal{I}}) e^{-i\tau H'}(\mathcal{I}\otimes|0\rangle)\| \\ &\quad\leq \int_0^{\delta t} \mathrm{d}\tau \Big[O(1)\|(\mathcal{I}\otimes\langle 0|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}} H' e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|0\rangle)\| \\ &\qquad + O((\delta t)^{1/2})\|(\mathcal{I}\otimes\langle 1|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}} H' e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|0\rangle)\| \\ &\qquad + O((\delta t)^{1/2})\|(\mathcal{I}\otimes\langle 0|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}} H' e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|1\rangle)\| \\ &\qquad + O(\delta t)\|(\mathcal{I}\otimes\langle 1|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}} H' e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|1\rangle)\| \Big] \\ &\quad= O((\delta t)^2) \ ,\end{aligned} \tag{3.252}$$

using (3.246a-3.246d). Similarly, we can bound

$$\begin{aligned}
&\|(\mathcal{I}\otimes\langle 1|)E(\mathcal{I}\otimes|0\rangle)\| \\
&\quad\le \int_0^{\delta t} \mathrm{d}\tau \|(\mathcal{I}\otimes\langle 1|)e^{-i(\delta t-\tau)(H_{\text{else}}\otimes\mathcal{I}+H')}e^{-i\tau H_{\text{else}}\otimes\mathcal{I}}(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}}H'e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})e^{-i\tau H'}(\mathcal{I}\otimes|0\rangle)\| \\
&\quad\le \int_0^{\delta t} \mathrm{d}\tau \Big[O((\delta t)^{1/2})\|(\mathcal{I}\otimes\langle 0|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}}H'e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|0\rangle)\| \\
&\qquad + O(1)\|(\mathcal{I}\otimes\langle 1|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}}H'e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|0\rangle)\| \\
&\qquad + O(\delta t)\|(\mathcal{I}\otimes\langle 0|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}}H'e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|1\rangle)\| \\
&\qquad + O((\delta t)^{1/2})\|(\mathcal{I}\otimes\langle 1|)(H' - e^{i\tau H_{\text{else}}\otimes\mathcal{I}}H'e^{-i\tau H_{\text{else}}\otimes\mathcal{I}})(\mathcal{I}\otimes|1\rangle)\| \Big] \\
&\quad= O((\delta t)^{3/2}) \ . \qquad (3.253)
\end{aligned}$$

With the bounds (3.252) and (3.253) on the blocks of the Trotter error we can now conclude that

$$\begin{aligned}
e^{-i\delta t(H'+H_{\text{else}}\otimes\mathcal{I})}(|\psi\rangle\otimes|0\rangle) &= e^{-i\delta t H_{\text{else}}\otimes\mathcal{I}}e^{-i\delta t H'}(|\psi\rangle\otimes|0\rangle) + E(|\psi\rangle\otimes|0\rangle) \\
&= e^{-i\delta t H_{\text{else}}\otimes\mathcal{I}}(e^{-i\delta t H}|\psi\rangle + O((\delta t)^2))\otimes|0\rangle + O((\delta t)^{3/2})\otimes|1\rangle \\
&\qquad + O((\delta t)^2)\otimes|0\rangle + O((\delta t)^{3/2})\otimes|1\rangle \\
&= \big(e^{-i\delta t H_{\text{else}}}e^{-i\delta t H}|\psi\rangle + O((\delta t)^2)\big)\otimes|0\rangle + O((\delta t)^{3/2})\otimes|1\rangle \ , \qquad (3.254)
\end{aligned}$$

where in the second inequality we invoke Proposition 3.4.14. 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 (3.239) with $t = \delta t$, $A = -iH_{\text{else}}$, $B = -iH$, we obtain

$$\begin{aligned}
&e^{-i\delta t H_{\text{else}}}e^{-i\delta t H} - e^{-i\delta t(H+H_{\text{else}})} \\
&\qquad = -i\int_0^{\delta t} \mathrm{d}\tau e^{-i(\delta t-\tau)(H+H_{\text{else}})}e^{-i\tau H_{\text{else}}}(H - e^{i\tau H_{\text{else}}}He^{-i\tau H_{\text{else}}})e^{-i\tau H} \ . \qquad (3.255)
\end{aligned}$$

So by Lemma 3.6.14 we have

$$e^{-i\delta t H_{\text{else}}}e^{-i\delta t H} - e^{-i\delta t(H+H_{\text{else}})} \le O((\delta t)^2) \qquad (3.256)$$

Combining (3.256) with (3.254) completes the proof. ∎

4 Analogue quantum simulation with polylogarithmic interaction strengths by extrapolating within phases of matter

The following section is a lightly edited version of the following article:

> [2]: Dylan Harley and Matthias Christandl. Analogue quantum simulation with polylogarithmic interaction strengths by extrapolating within phases of matter. *arXiv preprint arXiv:2605.11285*, 2026.

Abstract

Simple families of quantum Hamiltonians can simulate general many-body systems at arbitrary precision through the use of perturbative gadgets, however this generally requires interaction strengths spanning many orders of magnitude which scale polynomially in the system size and inverse precision, resulting in physically unrealisable systems. In this work, we show that for non-critical systems these required scalings can be exponentially reduced through classical post-processing, by simulating the model at smaller energy scales and extrapolating observables to the perturbative limit. In particular, we show that both local and extensive properties of thermal states with exponentially decaying correlations and ground states with a sufficiently stable gap can be simulated using gadgets whose interaction strengths scale only polylogarithmically in the inverse precision and the system size. As a key tool, we develop a generalised treatment of the local Schrieffer-Wolff transformation for geometrically quasi-local Hamiltonians over many energy scales, facilitating the analysis of perturbative gadget Hamiltonians without extensive global energy penalities, which may be of independent interest.

4.1 Introduction

4.1.1 Background and motivation

The simulation of quantum systems has long been recognised as an important use-case for quantum technologies [16], and remains a key focus for applications in both the near and far term [140, 31]. In contrast to the digital case, in which Hamiltonian dynamics are implemented fault-tolerantly by a universal gate-based quantum computer [43, 180], analogue quantum simulation involves encoding the behaviour of a Hamiltonian of interest into a tunable simulator [51], so that static and dynamic properties of the target system can be inferred from the native static and dynamic properties of the simulator. Compared to digital quantum simulation, this offers a potentially more experimentally tractable approach for near-term probes of many-body physics. Moreover, hybrid algorithms, using analogue simulation as a subroutine, may offer a path to reduce fault-tolerant circuit overhead for simulation tasks [181, 182, 183]; indeed digital-analogue hybrid simulations have already been demonstrated in practice [184]. Typically, we consider a situation where an experimenter has access to a tuneable but limited simulator, for instance with the ability to adjust the pairwise interaction strengths between sites whose geometry is fixed. Though these limitations appear to fundamentally restrict the regime of applicability for such a simulator, it turns out that even a very simple such model can be used to simulate arbitrary many-body Hamiltonians via perturbation theory [56].

Since their initial application in establishing the QMA-hardness of the 2-local Hamiltonian problem [57], perturbative simulations have become a widespread and powerful tool in the theory of Hamiltonian complexity and analogue simulation [72, 135, 86, 87, 56]. The basic idea is to take some target Hamiltonian and replace its interactions term-by-term with so-called perturbative gadgets, which each use an ancillary qubit to mediate interactions between several sites. In this way, the target Hamiltonian can be realised as the effective low-energy theory of the perturbative simulation Hamiltonian, which may belong to a simpler family. By iteratively applying such constructions, it is possible to reduce any many-body Hamiltonian to, for example, a 2-local Heisenberg model in the plane as in Ref. [56]. The drawback of perturbative gadgets is their scaling behaviour: in order to simulate a target Hamiltonian on n sites up to some error $\epsilon > 0$, perturbative gadgets typically contain individual interaction terms which scale as $\mathrm{poly}(n, \epsilon^{-1})$ in order to ensure accurate convergence of the perturbative series. This leads to simulator Hamiltonians whose interactions range over many orders of magnitude, and which are unrealisable in practice even for modest system sizes.

The prior work Ref. [135] has established that n-independent interaction strengths (scaling only as $\mathrm{poly}(\epsilon^{-1})$) are sufficient to simulate the ground state energy of a Hamiltonian up to extensive error $\sim n\epsilon$, but this result does not extend to any other properties of the ground state, and the polynomial scaling remains prohibitive if high accuracy is desired. In a previous work [1] we proved a no-go result showing that, for a broad class of modular simulation techniques, polynomial scaling is generally unavoidable.

Using polynomial extrapolation methods which have previously found application

in similar problems for quantum error mitigation [185, 186], tensor network representations [187], and recently for digital quantum simulation [47, 188], we show that in many cases these scalings can be exponentially reduced to $\mathrm{poly}\log(n\epsilon^{-1})$ through classical postprocessing. In particular, we show that properties of Gibbs states and ground states of perturbative simulator Hamiltonians can be estimated by measuring the system at several different values of the perturbative parameter, and extrapolating to the infinite interaction strength limit (in which the simulation is perfect); see Figure 4.1 for an illustration. These results require some additional structure: we assume that the system we are attempting to simulate is not close to a phase transition, to avoid attempted extrapolation through critical points. In particular, it is sufficient that the Gibbs state (respectively ground state) of interest has an exponential decay of correlations (respectively constant spectral gap) which is stable to small perturbations (see Conditions (I)-(III) below, and surrounding discussion).

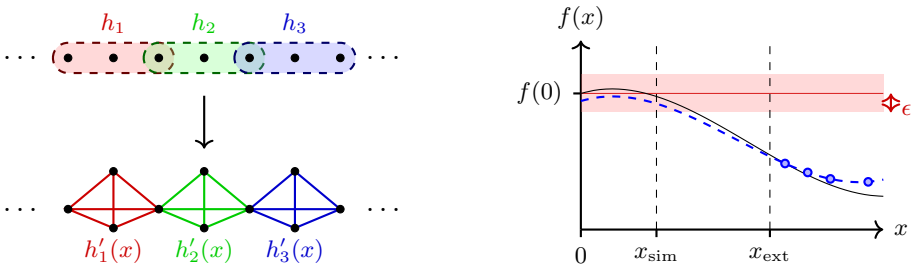
We expect this work may be of interest from several perspectives. Firstly, for practical simulations such as near-term approaches to Gibbs sampling, ground state preparation, and quantum phase estimation including analogue steps (recent proposals include e.g. Refs. [189, 183, 190, 191, 181, 182, 192]), our techniques open the door to using perturbative gadgets without prohibitive interaction scalings to access larger families of Hamiltonians. Secondly, from the perspective of Hamiltonian complexity theory, our work gives a many-to-one reduction between non-critical local Hamiltonian problems, and may be viewed as further evidence of the role of criticality in computational hardness (see Refs. [119, 120]). Additionally, for the theory of ground state and Gibbs state learning, the tools we develop may be used to design variations on the algorithms of e.g. Refs. [193, 194], which infer the values of observables of parametrised Hamiltonians from random samples. These prior works approximate local observables by linearly interpolating between locally similar sample points in parameter space; meanwhile our results offer accuracy guarantees for when higher-order functions can be fitted between data points.

4.1.2 Our contributions

In this section, we give a brief outline of the tools and contributions of the work; these come in three main parts. First, we prove extrapolation results for Hamiltonians under weak analytic perturbations (not in the simulation regime). Next, we develop tools for analysing (singular) gadget Hamiltonians using local Schrieffer-Wolff perturbation theory. Finally, combining these, we prove that properties of gadget Hamiltonians can themselves be extrapolated to the limit of perfect simulation. The structure of the main results in this work is illustrated in Figure 4.2.

Extrapolating within phases of matter

Initially, we consider a family of Hamiltonians $H(x)$ depending analytically on some parameter x (later, we will consider simulator Hamiltonians depending polynomially on x^{-1}). Letting $\rho(x)$ be the corresponding Gibbs or ground state of $H(x)$, and choosing



(a) (b)

Figure 4.1: (a) A 3-local Hamiltonian $H_{\text{tar}} = \sum_i h_i$ can be simulated by a 2-local simulator Hamiltonian $H'(x) = \sum_i h'_i(x)$ by replacing each term with a perturbative gadget containing interactions scaling polynomially with x^{-1} [57, 72]. (b) Behaviour of a simulated expectation value $f(x) := \text{tr}[O\rho'(x)]$ (black curve), for $\rho'(x)$ a ground or Gibbs state of the simulator Hamiltonian $H'(x)$. To ensure $|f(x) - f(0)| \leq \epsilon$, one must take $x \leq x_{\text{sim}} \sim 1/\text{poly}(n, \epsilon^{-1})$. We show that the same accuracy can be achieved by instead noisily sampling $f(x_k)$ at $x_k \sim x_{\text{ext}} = 1/\text{poly}\log(n\epsilon^{-1})$ (blue nodes) and extrapolating (blue curve) to $x = 0$.

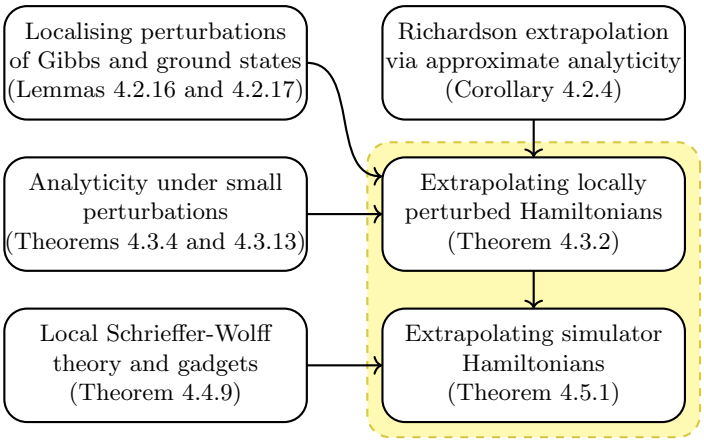


Figure 4.2: Structure of the key results of this work, with the main extrapolation results highlighted.

an observable O, we aim to estimate the value of a function $f(x) = \mathrm{tr}[O\rho(x)]$ at $x = 0$, from a set of m sample points $\{f(x_k)\}_{k=1}^m$ where $x_k > 0$ for all k. To this end, we use Richardson extrapolation [195, 196], which estimates the value of $f(0)$ by fitting a degree-$(m-1)$ polynomial to the samples $\{f(x_k)\}_{k=1}^m$, and simply evaluating this at $x = 0$.

There are two sources of error in Richardson extrapolation. Firstly, the choice of sample points $\{x_k\}_{k=1}^m$ may lead to an ill-conditioned problem to fit the polynomial, making the procedure very sensitive to errors. To overcome this issue, we follow Ref. [47] and choose so-called Chebyshev nodes, leading to an optimally well-conditioned problem. Secondly, we need to ensure that $f(x)$ has a good polynomial approximation; to this end, it is sufficient to show that $f(x)$ is well-approximated on the real line by a complex function $\tilde{f}(z)$, which is analytic and bounded in some region of interest.

Our problem thus reduces to showing that the functions of interest $f(x) = \mathrm{tr}[O\rho(x)]$ have sufficiently good analytic approximations. This is not immediately obvious: for Gibbs states $\rho_\beta(x) := e^{-\beta H(x)}/\mathrm{tr}[e^{-\beta H(x)}]$, extending the domain of x to the complex plane may lead to zeroes in the partition function $Z(x) := \mathrm{tr}[e^{-\beta H(x)}]$ in turn causing non-analyticities in $\rho_\beta(x)$. Though such behaviour cannot occur on the real axis, in the thermodynamic limit $n \to \infty$ the zeroes can approach the real axis as $\sim 1/n$, limiting the domain in which extrapolation is possible, see Figure 4.3 for an illustration. Meanwhile for ground states, $\rho(x)$ may not be continuous or even well-defined.

As a result, additional assumptions are necessary to guarantee that $f(x)$ can be approximated — physically speaking, this corresponds to assuming that we do not attempt to extrapolate through a phase transition. As discussed below, these assumptions are typically hard to prove rigorously for explicit examples of Hamiltonians (partially on fundamental complexity-theoretic grounds [123]) and are slightly weaker versions of those used in prior works on learning theory [198, 194, 193]. We informally describe these assumptions below, and first introduce some notation. We assume the Hamiltonian $H(x)$ acts on a set of sites Γ, where $|\Gamma| = n$, with some underlying metric $\mathrm{dist} : \Gamma \times \Gamma \to \mathbb{R}_{\geq 0}$. We assume that this describes a D-dimensional space (that is, balls of radius $r \geq 0$ with respect to dist have volume scaling as $\mathcal{O}(r^D)$). We decompose $H(x) = \sum_{A \subseteq \Gamma} H_A(x)$, where each $H_A(x)$ is supported only on the sites $A \subseteq \Gamma$. Our assumptions are then as follows (see Assumption 4.3.1):

(I) *Geometric quasi-locality:* We assume that $H(x)$ has exponentially decaying interactions, in the sense that for any sites $x, y \in \Gamma$, the total strength of interactions connecting them decays exponentially with distance:

$$\sum_{\substack{A \subseteq \Gamma \\ x,y \in A}} \|H_A(s)\| \leq \exp\left(-\Omega(\mathrm{dist}(x,y))\right) . \tag{4.1}$$

(II) *Uniform decay of correlations (for Gibbs states):* We assume that, for some region $x \in [0, x_*]$, $x_* > 0$, the Gibbs state $\rho_\beta(x) \propto e^{-\beta H(x)}$ has exponentially decaying correlations. For $\Gamma' \subseteq \Gamma$ and $x, y \in [0, x_*]$, we assume the same holds for the Gibbs

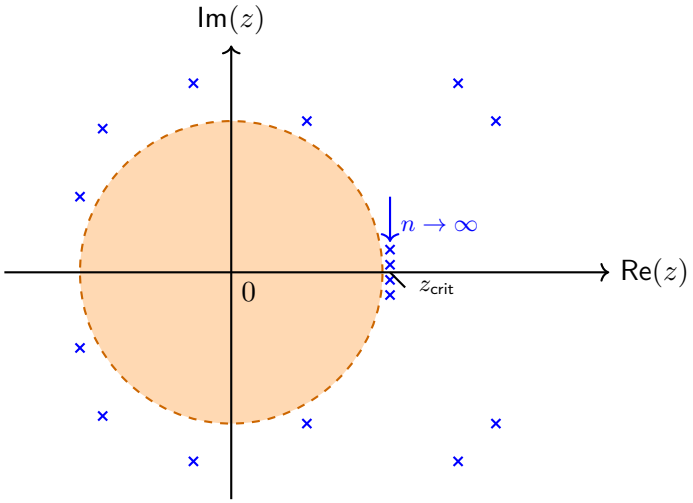


Figure 4.3: Zeroes (blue) of the complex partition function $Z(z) := \mathrm{tr}[e^{-\beta H(z)}]$, for an analytic family of Hamiltonians $z \mapsto H(z)$. Zeroes of $Z(z)$ correspond to non-analytic points of the Gibbs state $\rho_\beta(z) := e^{-\beta H(z)} / \mathrm{tr}[e^{-\beta H(z)}]$, so that the radius of convergence for a Taylor expansion for $\rho_\beta(z)$ is given by the distance to the nearest zero. In a region where $\rho_\beta(z)$ is analytic (shaded orange), Richardson extrapolation can be performed along the real axis. At phase transitions $z_{\text{crit}} \in \mathbb{R}$, the complex zeroes of $Z(z)$ "pinch" the real axis in the thermodynamic limit $n \to \infty$, preventing extrapolation past these points. A similar problem, where Z is instead viewed as a function of β, is studied in Ref [197].

states $\rho_\beta(x, y; \Gamma') \propto e^{-\beta H(x,y;\Gamma')}$, where $H(x, y; \Gamma')$ is the Hamiltonian where interactions acting on Γ' are given parameter x, whilst all others are given parameter y:

$$H(x, y; \Gamma') := \sum_{A \subseteq \Gamma'} H_A(x) + \sum_{\substack{A \subseteq \Gamma \\ A \cap (\Gamma \setminus \Gamma') \neq \emptyset}} H_A(y) \ . \tag{4.2}$$

(III) *Uniform spectral gap (for ground states):* We assume that, for some $x \in [0, x_*]$, the Hamiltonian $H(x)$ has a unique ground state with a constant spectral gap $\gamma > 0$. We assume the same holds for the Hamiltonians $H(x, y; \Gamma')$ as defined in Eq. (4.2).

Condition (I) ensures that the Hamiltonians $H(x)$ are sufficiently localised so that Lieb-Robinson bounds [54, 199] can be applied. Condition (II) (respectively (III)) ensures that the Gibbs (respectively ground) state is far from phase transitions. That this behaviour holds over a range of parameters of $H(x, y; \Gamma')$ is essential to argue that the x-dependence can be effectively ignored far away from a local observable of interest, as we describe below. Conditions (II)-(III) are similar to those used in Refs. [198, 194, 193] to obtain provable learning results for properties of quantum states within phases of matter. In particular, Condition (II) is satisfied automatically at sufficiently high temperature [200, 201], for translationally-invariant systems in one dimension [202, 203], or for commuting or one-dimensional systems above a thermal phase transition [197]. Condition (III) is implied by the stronger condition of local topological quantum order [204], which holds for toy models such as the toric code — however in general proving whether a spectral gap persists for a family of Hamiltonians in the $n \to \infty$ limit is computationally intractable [123].

Below we summarise the main result of this section, establishing that local properties of such systems can be extrapolated, and which is stated fully and proved in Section 4.3.

Result 1 (Extrapolation within phases of matter — see Theorem 4.3.2 and Corollary 4.3.3)**.** *Suppose $H(x)$ is a family of Hamiltonians depending analytically on x and satisfying Conditions* (I) *and* (II) *(respectively* (I) *and* (III)*), and let O_A be an observable supported on sites $A \subseteq \Gamma$, $|A| = \mathcal{O}(1)$. Let $f(x) = \mathrm{tr}[O_A \rho(x)]$, for $\rho(x)$ a constant-temperature Gibbs state (respectively the ground state) of $H(x)$. Then the value of $f(0)$ can be calculated up to any desired accuracy $\epsilon > 0$ via Richardson extrapolation from the values of $f(x_k)$ for $x_k \geq 1/\operatorname{poly}\log(\epsilon^{-1})$.*

The case of extensive observables $O = \sum_{A \subseteq \Gamma} O_A$ can also be dealt with by writing O as a sum of local observables and arguing that they can be individually extrapolated (see Appendix 4.6.5).

The proof of Result 1 requires two steps. Firstly, we argue that (up to a small error) the value of $f(x)$ only depends on the variation of $H(x)$ within a radius $r \sim \log \epsilon^{-1}$ of A — in other words, showing that we can assume that $\rho(x)$ is in fact the Gibbs (respectively ground) state of $H(x, 0; B_r(A))$, where $B_r(A)$ contains all sites within a radius r of A.

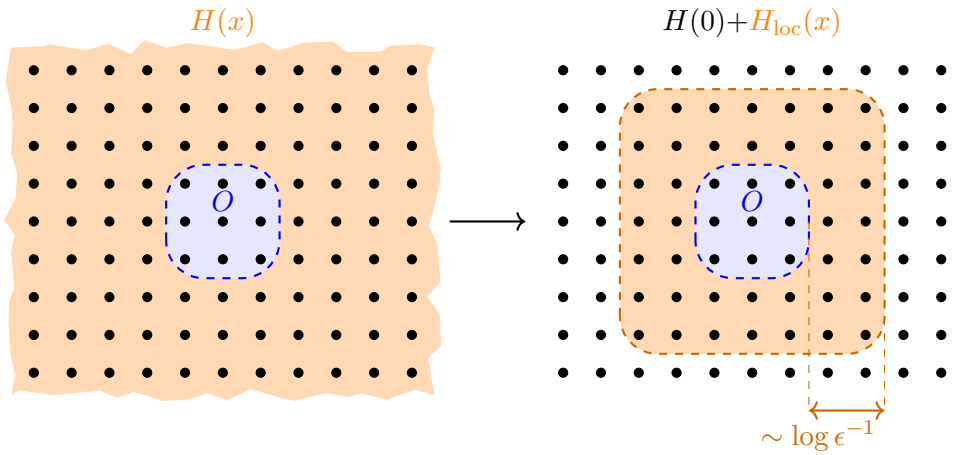


Figure 4.4: Away from criticality, the local expectation value $\mathrm{tr}[O\rho(x)]$ (for $\rho(x)$ a ground or Gibbs state) of a geometrically local Hamiltonian $H(x)$ can be approximated by ignoring the variation of the Hamiltonian outside of a region of $\log \epsilon^{-1}$ from the support of O (shaded in blue). Thus, for extrapolating properties of the perturbed Hamiltonian, it is sufficient to establish that $\mathrm{tr}[O\rho(x)]$ is well-approximated by an analytic function for Hamiltonian perturbations of magnitude $\|H_{\mathrm{loc}}(x)\| \sim \mathrm{poly} \log \epsilon^{-1}$.

See Figure 4.4 for an illustration. This step requires our assumptions (I)-(III), and uses the tools of quantum belief propagation [121] (respectively the spectral flow [122, 205]) to show that the variation of $\rho(x)$ with respect to x can be expressed as a local evolution, and the influence of distant interactions on O_A is controlled with Lieb-Robinson bounds [54, 199].

Having reduced to local perturbations supported on a small region around O_A, the second step of the proof involves establishing that $\rho(x)$ depends (approximately) analytically on x for sufficiently small perturbations of this form. For Gibbs states, this requires establishing lower bounds on the magnitude of the partition function $\mathrm{tr}[e^{-\beta(H(0)+V(x))}]$ for $\|V(x)\| = \mathcal{O}(1)$, to ensure that the Gibbs state has no non-analyticities. This follows from a series of arguments involving elementary linear algebra. For ground states, we write the slightly perturbed ground state in terms of a time-ordered exponential of the spectral flow operator, and establish that this is approximately analytic and bounded for small perturbations.

Local Schrieffer-Wolff perturbation theory

Note that Result 1 only applies to Hamiltonians depending analytically on the perturbation parameter x which we aim to send to 0. A priori, this appears to be a different regime to the case of perturbative simulator Hamiltonians, which typically depend polynomially on x^{-1} and are thus singular at $x = 0$. Our next results involve a systematic analysis of the properties of such Hamiltonians, ultimately aiming to circumvent the singularity by restricting to an effective Hamiltonian which depends only analytically on

x.

The general construction for a simulator Hamiltonian $H'(x)$ is as follows. The set of sites Γ is partitioned into $\Gamma = \Gamma_{\mathsf{eff}} \cup \Gamma_{\mathsf{anc}}$, where the sites $i \in \Gamma_{\mathsf{anc}}$ are ancillary sites used to mediate interactions and induce an effective Hamiltonian on Γ_{eff} (for simplicity we assume these are qubits, with local basis states $|0_i\rangle$ and $|1_i\rangle$). In particular, $H'(x)$ then takes the form

$$H'(x) = \Delta x^{-d} \sum_{i \in \Gamma_{\mathsf{anc}}} |1_i\rangle\langle 1_i| + \sum_{\alpha \geq 1} x^{\alpha - d} H^{(\alpha)} \ , \tag{4.3}$$

where $\Delta > 0$ and $|1_i\rangle\langle 1_i|$ is a 1-local projector onto acting on site i, and for some $d \geq 0$ referred to as the degree of the simulation. The x^{-d} term then incurs an energy penalty on all configurations of the ancillary sites except the all-zeroes state $|\mathbf{0}_{\mathsf{anc}}\rangle$. Meanwhile, the lower-order terms $H^{(\alpha)}$ are geometrically local Hamiltonians with interactions between Γ_{anc} and Γ_{eff}. For sufficiently small $x > 0$ (depending on n), the energy penalty from the x^{-d} term is large enough to induce a global energy gap, leading to the projector onto a low-energy space of $H'(x)$ given (approximately) by $\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}\rangle\langle\mathbf{0}|_{\mathsf{anc}}$. Restricting to this low-energy space, the Γ_{eff} system evolves under some different, effective Hamiltonian.

More concretely, $H'(x)$ is block-diagonal with respect to the projector $e^{-T(x)}(\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle\langle\mathbf{0}_{\mathsf{anc}}|)e^{T(x)}$, for $T(x)$ an anti-hermitian operator generating the small rotation $e^{T(x)}$. This allows us to define the effective low-energy Hamiltonian $H_{\mathsf{eff}}(x)$ as

$$H_{\mathsf{eff}}(x) := (\mathbb{1}_{\mathsf{eff}} \otimes \langle\mathbf{0}_{\mathsf{anc}}|)e^{T(x)} H'(x) e^{-T(x)} (\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle) \ . \tag{4.4}$$

Surprisingly, it is possible to construct such $T(x)$ and $H_{\mathsf{eff}}(x)$ even in the case when x is not small enough to induce a global energy gap: one can construct a power series $T(x) = \sum_{q \geq 1} x^q T^{(q)}$ (known as the local Schrieffer-Wolff transformation [206, 172]), such that $iT(x)$ is a bounded geometrically quasi-local Hamiltonian, and $e^{T(x)} H'(x) e^{-T(x)}$ is block-diagonal with respect to $\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle\langle\mathbf{0}_{\mathsf{anc}}|$. Moreover, the power series for $T(x)$ (and hence $H_{\mathsf{eff}}(x)$) is well-defined and convergent for x smaller than some constant $x_* > 0$, which is independent of the system size. This conclusion is in contrast to the usual (global) Schrieffer-Wolff transformation [172], for which x must be taken sufficiently small that $H'(x)$ has a global energy gap above the effective space $\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle\langle\mathbf{0}_{\mathsf{anc}}|$.

Our treatment of local Schrieffer-Wolff perturbation theory is similar to Ref. [172], with two major differences. Firstly, we explicitly consider Hamiltonians of the form Eq. (4.3) containing higher-order powers of x, allowing us to explicitly obtain $H_{\mathsf{eff}}(x)$ as a power series in x, in contrast to Ref. [172] in which the Hamiltonians involved have the form $H_0 + xV$. Secondly, our analysis makes use of local Hamiltonian norms introduced in Ref. [199], which naturally allow the application of Lieb-Robinson bounds. In fact, using these norms turns out to make the analysis slightly simpler, leading to provably convergent power series for $T(x)$ and $H_{\mathsf{eff}}(x)$ which are instead truncated at finite order in Ref. [172] — however we in turn require the assumption of geometric locality, which is not present in Ref. [172].

If $H_{\mathsf{eff}}(x)$ is analytic in x (i.e., the negative powers in its series expansion all cancel), we say that $H'(x)$ is a simulator Hamiltonian, and we thus recover a well-defined effective

Hamiltonian in the limit $x \to 0$. We formalise this notion in Section 4.4.3, and show that it recovers familiar notions of simulation formalised in Refs. [87, 56, 1]. We also prove a general result (see Theorem 4.6.13) establishing that such Hamiltonians can be built from "gadgets" used in parallel: roughly speaking, given several $\{h'_j(x)\}_j$ such that each $h'_j(x)$ is a degree $d \leq 3$ simulation with effective Hamiltonian $h_{\mathsf{eff},j}(x)$, the combined Hamiltonian $H'(x) = \sum_j h'_j(x)$ yields an effective Hamiltonian $H_{\mathsf{eff}}(x) = \sum_j h_{\mathsf{eff},j}(x)$.

Our treatment of local Schrieffer-Wolff perturbation theory allows us to prove several useful properties of simulator Hamiltonians, which we summarise below as Result 2. To our knowledge, this is the first formal treatment of Hamiltonian gadgets for quantum simulation with low energies via the local Schrieffer-Wolff transformation, and we expect these tools may be independently useful.

Result 2 (Properties of simulator Hamiltonians — see Theorem 4.4.9)**.** *Let $H'(x)$ be a degree-d simulator Hamiltonian constructed as in Eq.* (4.3)*, with analytic effective Hamiltonian $H_{\mathsf{eff}}(x)$. There is a constant $x_* > 0$ such that, for $x \leq x_*$, the following holds:*

(I) Effective Hamiltonian: *The power series for $H_{\mathsf{eff}}(x)$ converges, and $H_{\mathsf{eff}}(x)$ inherits the same quasi-locality properties as $H'(x)$ (in particular, if each $H^{(\alpha)}$ has exponentially decaying interactions, then so does $H_{\mathsf{eff}}(x)$).*

(II) Gibbs states: *Let $\rho'_\beta(x)$ and $\rho_{\beta,\mathsf{eff}}(x)$ be the Gibbs states of $H'(x)$ and $H_{\mathsf{eff}}(x)$ respectively at inverse temperature $\beta > 0$. Then*

$$\rho'_\beta(x) \approx e^{-T(x)}(\rho_{\beta,\mathsf{eff}}(x) \otimes |\mathbf{0}_{\mathsf{anc}}\rangle\langle\mathbf{0}_{\mathsf{anc}}|)e^{T(x)} \ , \tag{4.5}$$

where "$\approx$" denotes approximation in the trace norm, up to an error which decays exponentially with x^{-d}.

(III) Ground states: *Let $P'(x)$ and $P_{\mathsf{eff}}(x)$ be the ground state projectors of $H'(x)$ and $H_{\mathsf{eff}}(x)$ respectively. Then*

$$P'(x) = e^{-T(x)}(P_{\mathsf{eff}}(x) \otimes |\mathbf{0}_{\mathsf{anc}}\rangle\langle\mathbf{0}_{\mathsf{anc}}|)e^{T(x)} \ . \tag{4.6}$$

Extrapolation of simulator Hamiltonians

The third main result of this work can be viewed as a combination of Results 1 and 2; it establishes that properties of simulator Hamiltonians (of the form given in Eq. (4.3)) can be extrapolated to the limit $x \to 0$ corresponding to perfect simulation. We state this result below; see Figure 4.5 for a summary of the interaction strengths necessary to compute different quantities.

Result 3 (Extrapolation of simulator Hamiltonians — see Theorem 4.5.1 and Corollary 4.5.2)**.** *Suppose $H'(x)$ is a family of simulator Hamiltonians as in Eq.* (4.3)*, such that the effective Hamiltonian $H_{\mathsf{eff}}(x)$ satisfies the conditions of Result 1. Let O_A be an observable supported on sites $A \subseteq \Gamma_{\mathsf{eff}}$, $|A| = \mathcal{O}(1)$. Let $\rho'(x)$ and $\rho_{\mathsf{eff}}(x)$ be the*

Simulation task	**State**	**Interaction strengths**
Full-spectrum simulation	Gibbs	$\mathrm{poly}(n, \epsilon^{-1})$
(See Refs. [87, 56])	Ground	$\mathrm{poly}(n, \epsilon^{-1})$
Ground state energy simulation [135]	Ground	$\mathrm{poly}(n\epsilon^{-1})$
Local observable extrapolation	Gibbs	$\mathrm{poly}\log(n\epsilon^{-1})$
(See Theorem 4.5.1, Corollary 4.5.2)	Ground	$\mathrm{poly}\log(\epsilon^{-1})$
Extensive observable extrapolation	Gibbs	$\mathrm{poly}\log(n\epsilon^{-1})$
(See Theorem 4.6.20, Corollary 4.6.21)	Ground	$\mathrm{poly}\log(n\epsilon^{-1})$

Figure 4.5: Comparison of the interaction strengths required in a simulator Hamiltonian $H'(x)$ to estimate properties of a geometrically local target Hamiltonian H_{tar} up to (additive) error $\epsilon > 0$. In particular, we aim to estimate quantities of the form $\mathrm{tr}[O\rho]$, where O is an observable and ρ is a ground state or Gibbs state of H_{tar}, by measuring corresponding observables on $H'(s)$. We separately consider local observables, for which O is supported on a constant number of sites, and extensive observables where O is allowed to contain quasi-local interactions throughout the system (for example, $O = H_{\mathrm{tar}}$). Full spectrum simulation as in Refs. [87, 56] captures all properties of the target Hamiltonian to any desired accuracy without any need for extrapolation, and works in full generality. Via extrapolation we establish that the same techniques can be used with exponentially weaker interaction strengths, for systems away from phase transitions.

constant-temperature Gibbs states (respectively the ground states) of $H'(x)$ and $H_{\mathsf{eff}}(x)$, and define $f'(x) := \mathrm{tr}[(O_A \otimes \mathbb{1}_{\mathsf{anc}})\rho'(x)]$. Then the value of $\mathrm{tr}[O_A\rho_{\mathsf{eff}}(0)]$ can be calculated up to any desired accuracy $\epsilon > 0$ via Richardson extrapolation, from values of $f'(x_k)$ with $x_k \geq 1/\mathrm{poly}\log(n\epsilon^{-1})$ (respectively $x_k \geq 1/\mathrm{poly}\log(\epsilon^{-1})$). In particular, the Hamiltonians $H'(x_k)$ contain interaction strengths of order $\mathrm{poly}\log(n\epsilon^{-1})$ (respectively $\mathrm{poly}\log(\epsilon^{-1})$).

The corresponding result for extensive quantities $O = \sum_{A\subseteq\Gamma} O_A$ (see Appendix 4.6.5) shows that interaction strengths of order $\mathrm{poly}\log(n\epsilon^{-1})$ are sufficient for both Gibbs states and ground states. In particular, choosing $O = H_{\mathsf{eff}}(0)$, the ground state energy for a non-critical system can be estimated to extensive precision $\sim \epsilon n$ using gadget interaction strengths of order $\sim \mathrm{poly}\log(\epsilon^{-1})$, as opposed to $\sim \mathrm{poly}(\epsilon^{-1})$ as in Ref. [135]. Our gadget formalism is sufficiently general to capture all mediator gadgets that we are aware of (such as Refs. [72, 87, 56]). As a concrete example, this includes reductions to simulate 3-local Hamiltonians with only 2-local interactions, as we describe in Section 4.5.3 — however our results are much more general and apply to all mediator gadgets in the literature which we are aware of. We note however that geometric locality is essential, so our results do not apply to simulations which radically alter the global geometry.

4.1.3 Discussion and future directions

In this work, we have proved how the overhead for analogue quantum simulation (as measured by local interaction strengths) can be exponentially reduced from $\mathrm{poly}(n, \epsilon^{-1})$ to $\mathrm{poly}\log(n\epsilon^{-1})$, for Gibbs state and ground state properties in non-critical systems on n sites up to precision ϵ. To this end, we used tools from many-body theory to extend previous works on quantum learning [198, 194, 193], establishing that such properties vary approximately analytically for analytically perturbed Hamiltonians. This allows the application of Richardson extrapolation [195] for polynomial interpolation. Moreover, we have analysed the behaviour of perturbative gadget Hamiltonians in the low-energy regime (that is, without a global energy gap) using the local Schrieffer-Wolff transformation [172], in particular proving that Gibbs and ground state properties are faithfully simulated up to high precision. This gives a clean reduction from singular simulator Hamiltonians to analytic effective Hamiltonians for quantities of interest. As an example, our results can be applied for the simulation of 3-local Hamiltonians by 2-local simulators, though our treatment of perturbative simulation is general enough to encompass all other mediator gadget constructions we are aware of.

Non-criticality assumptions. One drawback of our results is that the non-criticality conditions we assume are difficult to prove (and even potentially undecidable [123]) for families of Hamiltonians other than simple toy models. This problem is also present in the related works [198, 193, 194], which use similar techniques to ours to localise Hamiltonian perturbations and infer local properties. Whilst it seems unlikely that this problem can be solved entirely, as phase transitions pose a fundamental barrier for extrapolation, it may be possible to establish weaker assumptions which are still sufficient. A potential extension in this direction comes from Ref. [207], which establishes learning results under the phase definition introduced by Ref. [208], wherein two states are said to lie in the same phase if they can be related by a short-time dissipative evolution. Furthermore, the numerics of Ref. [198] suggest that predicting properties within phases may work in practice in many situations where the required assumptions cannot be rigorously established in theory; it would be interesting to see whether similar empirical conclusions can be obtained in our case.

Applications for simulating dynamics. Our results apply to static properties of simulator Hamiltonians, that is, properties of their thermal and ground states. It is natural to ask whether the local Schrieffer-Wolff and classical post-processing techniques might also be useful for analysing the dynamics of such systems. Further work in this direction could build on recent works (see for example Refs. [52, 53]) analysing and bounding the propagation of errors in noisy analogue simulators, or using perturbative methods to encode the target dynamics into a space protected by a quantum error-detecting code [209].

Analyticity and phase transitions. One more fundamental mathematical question raised by this work involves the relationship between phase transitions and zeroes of the partition function. In particular, given a family of Hamiltonians $H(x)$ depending analytically on x, the Gibbs states $e^{-\beta H(x)}/\operatorname{tr}[e^{-\beta H(x)}]$ are analytic when extended to complex x except at zeroes of the partition function $Z(x) := \operatorname{tr}[e^{-\beta x}]$. Distance from the nearest zeroes of $Z(x)$ controls the radius of convergence of the Taylor series, and thus whether Richardson extrapolation can be directly applied (see Figure 4.3). In our proof of Result 1, we circumvent this issue using the assumption of exponential correlation decay, which allows us to approximately restrict the variation of $H(x)$ and bound the zeroes of $Z(x)$ away from $x = 0$. Nonetheless, it is an interesting theoretical question whether our assumption (II) is sufficient to directly prove exact analyticity of the Gibbs state, for local variations independent of the system size. Previous work [197] has examined a closely related problem, in which the partition function is viewed as a complex function of the inverse temperature β with the Hamiltonian fixed. In particular, the authors show that under some additional assumptions, an absence of zeroes near the real line of the partition function implies that distant correlations are exponentially small. In the classical case these conditions are known to be equivalent [210, 197], and it remains an open question to establish an analogous quantum result.

4.2 Preliminaries

4.2.1 Richardson extrapolation

On the classical postprocessing side, we will use Richardson extrapolation [195, 196], a method which has previously been applied in the different but related context of digital quantum simulation via Trotterisation [47, 188]. In this section we sketch the main ideas of this approach, and state the results necessary for our purposes.

We aim to compute the value of $f(0)$, for some unknown function $f : \mathbb{C} \to \mathbb{C}$ which we can only access at limited precision for inputs $x \in (0, x_{\max}]$. That is, we are given some set of estimates $\{\hat{f}(x_1), \ldots, \hat{f}(x_m)\}$ such that $|\hat{f}(x_k) - f(x_k)| \leq \delta$ for all k, for some error rate $\delta > 0$. Sampling from values of x closer to zero will be more expensive (in our case, this will ultimately correspond to implementing a simulator Hamiltonian with interaction strengths scaling as $\sim 1/\operatorname{poly}(x)$), so ideally we would like an extrapolation scheme such that $\min_k |x_k|$ need not be too small.

The strategy is as follows: we approximate $f(x)$ with a degree-$(m-1)$ polynomial $f(x) = c_0 + c_1 x + \cdots + c_{m-1}x^{m-1} + E_m(x)$ (up to some higher-order errors $E_m(x) = \mathcal{O}(x^m)$), and estimate the coefficients c_k by some $\hat{c}_k$, calculated by inverting the resulting set of linear equations

$$\hat{f}(x_k) = \hat{c}_0 + \hat{c}_1 x_k + \cdots + \hat{c}_{m-1} x_k^{m-1} \ , \quad \text{for } k = 1, \ldots, m \ . \tag{4.7}$$

The estimated coefficient $\hat{c}_0$ is then our approximation for $f(0)$. Notice that there are two sources of error in this procedure: the error $E_m(x)$ incurred by our polynomial approximation, and the sampling errors $\hat{f}(x_k) - f(x_k)$.

The former error can be controlled by bounding the Taylor series truncation error of the function $f(x)$, expanded around $x = 0$. To this end, we can use the following corollary of Cauchy's integral theorem:

Lemma 4.2.1 (Taylor series truncation error — see [197], Proposition 18)**.** *Suppose $f : \mathbb{C} \to \mathbb{C}$ is an analytic function bounded as $|f(x)| \leq M$ for $|x| \leq b$, where $b > 1$. Then the error of truncating $f(x)$ by a Taylor series of degree m in $|x| \leq 1$, $f(x) = c_0 + c_1 x + \cdots + c_m x^m + E_{m+1}(x)$, is bounded by*

$$\big|E_{m+1}(x)\big| \leq \frac{M}{b^m(b-1)} \ , \quad \textit{for } |x| \leq 1 \ . \tag{4.8}$$

The latter error is slightly more subtle; it depends on the conditioning of the linear system defined by Eq. (4.7), which is in turn very sensitive to the choice of sampling points x_j. The linear system corresponds to a Vandermonde matrix, which we can solve explicitly for $\hat{c}_0$ as

$$\begin{aligned}
\hat{c}_0 &= \begin{pmatrix} 1 & 0 \ldots & 0 \end{pmatrix} \begin{pmatrix} 1 & x_1 & \ldots & x_1^{m-1} \\ 1 & x_2 & \ldots & x_2^{m-1} \\ \vdots & & \ddots & \vdots \\ 1 & x_m & \ldots & x_m^{m-1} \end{pmatrix}^{-1} \begin{pmatrix} \hat{f}(x_1) \\ \hat{f}(x_2) \\ \ldots \\ \hat{f}(x_m) \end{pmatrix} \\
&= \sum_k \Big(\prod_{j \neq k} \frac{x_j}{x_j - x_k} \Big) \hat{f}(x_k) \ .
\end{aligned} \tag{4.9}$$

Meanwhile, the true value of $f(0) = c_0$ is given by

$$c_0 = \begin{pmatrix} 1 & 0 & \ldots & 0 \end{pmatrix} \begin{pmatrix} 1 & x_1 & \ldots & x_1^{m-1} \\ 1 & x_2 & \ldots & x_2^{m-1} \\ \vdots & & \ddots & \vdots \\ 1 & x_m & \ldots & x_m^{m-1} \end{pmatrix}^{-1} \begin{pmatrix} f(x_1) - E_m(x_1) \\ f(x_2) - E_m(x_2) \\ \ldots \\ f(x_m) - E_m(x_m) \end{pmatrix} \tag{4.10}$$

$$= \sum_k \Big(\prod_{j \neq k} \frac{x_j}{x_j - x_k} \Big) \big(f(x_k) - E_m(x_k) \big) \ , \tag{4.11}$$

and hence the error in the extrapolated value can be bounded by

$$\begin{aligned}
|c_0 - \hat{c}_0| &= \bigg| \sum_k \Big(\prod_{j \neq k} \frac{x_j}{x_j - x_k} \Big) \big(f(x_k) - \hat{f}(x_k) - E_m(x_k) \big) \bigg| \\
&\leq \big(\max_k |f(x_k) - \hat{f}(x_k)| + \max_k |E_m(x_k)| \big) \sum_k \prod_{j \neq k} \bigg| \frac{x_j}{x_j - x_k} \bigg| \\
&= \Big(\delta + \sup_{|z| \leq x_{\max}} |E_m(z)| \Big) \alpha(x_1, \ldots, x_m) \ ,
\end{aligned} \tag{4.12}$$

where we have used our assumption that $|f(x_k) - \hat{f}(x_k)| \leq \delta$, and defined the dimensionless condition number

$$\alpha(\mathbf{x}) = \sum_k \prod_{j \neq k} \left| \frac{x_j}{x_j - x_k} \right| . \tag{4.13}$$

A naive choice of sampling points $\mathbf{x}$, such as uniformly spaced $x_k = x_{\max} k/m$, will lead to poor conditioning $\alpha(\mathbf{x}) = e^{\Omega(m)}$, making the computation extremely sensitive to noise. Better choices exist, however: in particular the Chebyshev nodes defined by

$$x_k = x_{\max} \sin^2 \left(\frac{(2k-1)\pi}{4m} \right) , \tag{4.14}$$

lead to a scaling of $\alpha(\mathbf{x}) = \mathcal{O}(\log m)$ (see Ref [47]). In Appendix 4.6.1 we give an elementary proof of this fact, with the concrete upper bound $\alpha(\mathbf{x}) \leq 3\log(m)$ for $m \geq 2$. Combining the bound Eq. (4.12) with the particular choice of Chebyshev nodes and Lemma 4.2.1, we arrive at the following general extrapolation result:

Theorem 4.2.2 (Richardson extrapolation with Taylor series truncation)**.** *Let $f : \mathbb{C} \to \mathbb{C}$ be an analytic function which is bounded as $|f(x)| \leq M$ whenever $x \in \mathbb{C}$ satisfies $|x| \leq R$. Using the Chebyshev nodes*

$$x_k = \frac{R}{2} \sin^2 \left(\frac{(2k-1)\pi}{4m} \right) \quad \text{for } k = 1, \ldots, m , \tag{4.15}$$

assume we have noisy samples $\hat{f}(x_k)$ such that $|\hat{f}(x_k) - f(x_k)| \leq \delta$ for all k, and compute $\hat{c}_0$ using Eq. (4.9)*. Then the extrapolation error from the true value of $f(0)$ is bounded by*

$$\left| f(0) - \hat{c}_0 \right| \leq \left(\delta + 2^{-m} M \right) \cdot 3\log(m) . \tag{4.16}$$

In particular, in order to bound the right-hand side of Eq. (4.16) below some $\epsilon > 0$, it is sufficient to take $m = \mathcal{O}(\log M\epsilon^{-1})$ samples at precision $\delta = \epsilon / \mathcal{O}(\log\log \epsilon^{-1})$. This leads to

$$x_{\min} = \min_k |x_k| = \frac{R}{\operatorname{poly}\log(M\epsilon^{-1})} . \tag{4.17}$$

Analytic approximations

In this work, we will often be in the situation where the function f we wish to extrapolate is not analytic, and hence Theorem 4.2.2 cannot immediately be applied. Instead, our strategy will be to show that f can be well-approximated by a different, genuinely analytic, function $\tilde{f}$, and to then apply Theorem 4.2.2 to this function. It will be useful to introduce the following definition.

Definition 4.2.3 (Analytic approximation)**.** Let $f : \mathbb{R} \to \mathbb{C}$ be a function, and let $\delta, M, R \geq 0$. We say that f has a (δ, M, R)-analytic approximation if there exists an analytic function $\tilde{f} : \mathbb{C} \to \mathbb{C}$ such that

$$\sup_{x \in [0,R]} |f(x) - \tilde{f}(x)| \leq \delta \ , \quad \sup_{|z| \leq R} |\tilde{f}(z)| \leq M \ . \tag{4.18}$$

Corollary 4.2.4 (Richardson extrapolation with approximate analyticity)**.** *Let $f : \mathbb{R} \to \mathbb{C}$ be a function. Assume that f has a (δ, M, R)-analytic approximation. Then, for any $m \geq 1$, an estimate $\hat{c}_0$ of the value of $f(0)$ can be extrapolated from the Chebyshev nodes $\{x_k\}_{k=1}^m$ as in Theorem 4.2.2 with sampling error δ, with error*

$$|f(0) - \hat{c}_0| \leq (\delta + 2^{-m} M)\, \mathcal{O}(\log m) \ , \tag{4.19}$$

and where where $x_{\min} := \min_k x_k$ is bounded as

$$x_{\min} = \Theta(R/m^2) \ . \tag{4.20}$$

Proof of Corollary 4.2.4. We view the samples of f as noisy samples of $\tilde{f}$, where the error is bounded by 2δ by the triangle inequality. Applying Theorem 4.2.2, this yields an estimate $\hat{c}_0$ satisfying

$$|\tilde{f}(0) - \hat{c}_0| \leq (2\delta + 2^{-m} M)\, \mathcal{O}(\log m) \ . \tag{4.21}$$

The result then follows using the fact that $|f(0) - \tilde{f}(0)| \leq \delta$, and absorbing the constant factor into the $\mathcal{O}(\log m)$. ∎

4.2.2 Many-body states

Local Hamiltonians and F-norms

In this section we will establish some notation and technical tools for the treatment of geometrically local Hamiltonians. We focus on many-body Hamiltonians on systems of n sites, with local Hilbert spaces denoted by $\mathcal{H}_i$, $i \in \Gamma$, $|\Gamma| = n$; usually we will think of qubits $\mathcal{H}_i \cong \mathbb{C}^2$, but this restriction is not necessary. The full Hilbert space is then denoted by $\mathcal{H} := \bigotimes_{i \in \Gamma} \mathcal{H}_i$. Any Hamiltonian $H \in \mathrm{Herm}(\mathcal{H})$ can be decomposed into local terms of the form

$$H = \sum_{A \subseteq \Gamma} H_A \ , \tag{4.22}$$

where each term H_A acts non-trivially only on the sites $A \subseteq \Gamma$, and acts as a tensor product of identity operators on all other sites. Though such a decomposition is generally not unique, there exists a canonical choice: write H as a sum over tensor products of Pauli matrices (which form a basis for $\mathrm{Herm}(\mathcal{H})$) and choose H_A to be the sum of those acting only on A. The Hamiltonian H is said to be *k-local* if $H_A = 0$ whenever $|A| > k$.

Following Ref. [199], we are interested in geometrically local Hamiltonians which may not be strictly k-local but with interactions whose strength generally decays over long distances according to some function F. We therefore assume the structure of a metric on Γ denoted by $\text{dist} : \Gamma \times \Gamma \to [0, \infty)$ encoding the physical positions of the n sites. The ball of radius r centred at $i \in \Gamma$ is defined as

$$B_r(i) := \{j \in \Gamma : \text{dist}(i, j) \leq r\} \ . \tag{4.23}$$

We then say that Γ has dimension $D \geq 1$ if there exists a constant $k_D > 0$ independent of n such that

$$\sup_{i \in \Gamma, r \geq 1} |B_r(i)| r^{-D} \leq k_D \ . \tag{4.24}$$

Given a non-increasing function $F : [0, \infty) \to [0, \infty)$, the $\|\cdot\|_F$-norm of a Hamiltonian is defined as

$$\|H\|_F := \sup_{i,j \in \Gamma} \frac{1}{F(\text{dist}(i, j))} \sum_{\substack{A \subseteq \Gamma \\ i,j \in A}} \|H_A\| \ . \tag{4.25}$$

Intuitively, this definition ensures that the total strength of interactions coupling sites $i, j \in \Gamma$ decays as $F(\text{dist}(i, j))$, since

$$\sum_{\substack{A \subseteq \Gamma \\ i,j \in A}} \|H_A\| \leq F(\text{dist}(i, j)) \|H\|_F \ . \tag{4.26}$$

Physical states and criticality

Given a family of Hamiltonians $\{H(x)\}_{x \in \mathcal{X}}$ (for some possibly multidimensional parameter space $\mathcal{X}$), we will typically denote the ground state of $H(x)$ (when this is well-defined) by $|\psi_0(x)\rangle$, and the Gibbs state at inverse temperature $\beta \geq 0$ by $\rho_\beta(x) := e^{-\beta H(x)}/Z_\beta(x)$, where $Z_\beta(x) = \text{tr}[e^{-\beta H(x)}]$ is the partition function. We can generally only expect extrapolation of physical properties to be reliable away from phase transitions (see e.g. Ref. [197] for a different connection between analyticity and correlation decay in thermal states). For ground states, we characterise criticality in terms of the spectral gap; ground state $|\psi_0(x)\rangle$ does not undergo a phase transition if the Hamiltonians $H(x)$ have a constant gap between the ground and first excited energies. We summarise this with the following definition.

Definition 4.2.5 (Uniformly gapped ground states)**.** A family of Hamiltonians $\{H(x)\}_{x \in \mathcal{X}}$ is uniformly gapped with gap $\gamma > 0$ if, for all $x \in \mathcal{X}$, the ground space of $H(x)$ is one-dimensional and its two lowest eigenvalues $\lambda_0(x) < \lambda_1(x)$ are separated by $\lambda_1(x) - \lambda_0(x) > \gamma$.

The assumption that the ground space is one-dimensional is necessary to extrapolate ground state properties, however it is not necessary if one aims to extrapolate the ground state energy (as $\lambda_0(x)$ may be is well-defined even when $|\psi_0(x)\rangle$ is not). For Gibbs states, we characterise non-criticality in terms of correlation decay:

Definition 4.2.6 (Exponential decay of correlations)**.** A state ρ on $\otimes_{x\in\Gamma}\mathcal{H}_x$ is said to satisfy an exponential decay of correlations with parameters $K, \xi > 0$ if, for all observables M_A and N_B supported on disjoint subsystems $A, B \subseteq \Gamma$, we have

$$\frac{|\operatorname{tr}[\rho(M_A \otimes N_B)] - \operatorname{tr}[\rho M_A]\operatorname{tr}[\rho N_B]|}{\|M_A\|\|N_B\|} \leq K|A||B|e^{-\operatorname{dist}(A,B)/\xi} \ . \tag{4.27}$$

Often (see e.g. Ref. [200]), a stronger definition is used with the sizes of the boundaries $|\partial A|$ and $|\partial B|$ in place of $|A|$ and $|B|$, but this is not necessary for our purposes as we will typically use this assumption in situations when one or both of $|A|$ and $|B|$ is constant. Note that gapped ground states exhibit exponential decay of correlations as in Definition 4.2.6 by Ref. [116], as do Gibbs states at sufficiently high temperature [200, 201]. The following definition, analogously to Definition 4.2.5, extends this assumption to hold uniformly over a family of states.

Definition 4.2.7 (Uniform exponential decay of correlations)**.** A family of states $\{\rho(x)\}_{x\in\mathcal{X}}$ is said to satisfy a uniform exponential decay of correlations with parameters $K, \xi > 0$ if $\rho(x)$ has exponentially decaying correlations as in Definition 4.2.6 for all $x \in \mathcal{X}$.

Quantum belief propagation and the spectral flow

Consider the Gibbs (respectively gapped ground) states of a family of Hamiltonians $\{H(x)\}_{x\in\mathcal{X}}$ on the sites Γ, whose local terms vary smoothly with x. It turns out that, under such variation, the Gibbs (gapped ground states) change in a predictable manner described by the framework of quantum belief propagation [121, 211, 212, 193] (respectively the spectral flow [122, 205, 199]). In both cases, the change in the Gibbs (gapped ground) state is controlled by an operator obtained by locally transforming $\partial_x H(x)$. We begin by stating the result for Gibbs states:

Lemma 4.2.8 (Quantum belief propagation [121, 211])**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ be a family of Hamiltonians which smoothly depends on x on the interval $\mathcal{X} \subseteq \mathbb{R}$. Then the derivative of the Gibbs state $\rho_\beta(x) = e^{-\beta H(x)}/\operatorname{tr}[e^{-\beta H(x)}]$ is given by*

$$\frac{\mathrm{d}}{\mathrm{d}x}\rho_\beta(x) = -\frac{1}{2}\beta\left\{\rho_\beta(x), \Phi_{H(x)}(\partial_x H(x))\right\} + \beta\rho_\beta(x)\operatorname{tr}\left[\rho_\beta(x)\Phi_{H(x)}(\partial_x H(x))\right] \ , \tag{4.28}$$

where $\Phi_H(\cdot)$ is the quantum belief propagation operator defined by

$$\Phi_H(X) := \int_{-\infty}^{\infty} \mathrm{d}t\kappa_\beta(t)e^{-itH}Xe^{itH} \ , \tag{4.29}$$

and $\kappa_\beta(t)$ is a function which decays exponentially away from $t = 0$, explicitly given by (see Ref. [212])

$$\kappa_\beta(t) = \frac{2}{\pi\beta}\log\frac{e^{\pi|t|/\beta}+1}{e^{\pi|t|/\beta}-1} \leq \frac{4}{\pi\beta}\cdot\frac{1}{e^{\pi|t|/\beta}-1} \ . \tag{4.30}$$

Proof of Lemma 4.2.8. An elementary proof of this fact is given, for example, in Ref. [200] Appendix B. The statement there only covers the case where $H(x)$ is linear, $H(x) = H(0) + xA$, but this is not a necessary restriction: the first step in the proof involves using Duhamel's identity to write

$$\frac{\mathrm{d}}{\mathrm{d}x} e^{-\beta H(x)} = -\beta \int_0^1 \mathrm{d}t e^{-\beta t H(x)} \partial_x H(x) e^{-\beta(1-t)H(x)} \ . \tag{4.31}$$

This is valid for any smooth function $H(x)$, and the remainder of the proof follows unchanged with $\partial_x H(x)$ in place of A. ∎

It is an immediate consequence (see Ref. [193]) of Eq. (4.28) that, for any observable O with expectation value $f_\beta(x) := \mathrm{tr}[O\rho_\beta(x)]$, the derivative of $f_\beta(x)$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}x} f_\beta(x) = -\beta \, \mathrm{Cov}_{\rho_\beta(x)} \left(O, \Phi_{H(x)}(\partial_x H(x)) \right) = -\beta \, \mathrm{Cov}_{\rho_\beta(x)} \left(\Phi_{H(x)}(O), \partial_x H(x) \right) \ , \tag{4.32}$$

where $\mathrm{Cov}_\rho(X, Y)$ is the operator covariance

$$\mathrm{Cov}_\rho(X, Y) := \frac{1}{2} \mathrm{tr}[\rho\{X, Y\}] - \mathrm{tr}[\rho X] \, \mathrm{tr}[\rho Y] \ . \tag{4.33}$$

Since uniform correlation decay as in Definition 4.2.7 gives an upper bound on the covariance between spatially separated observables, this form can be leveraged to localise the dependence of $f_\beta(x)$ on $\partial_x H(x)$ to terms in a small region around O (we formalise this fact in Lemma 4.2.16).

For ground states, we have the following qualitatively similar (though technically distinct) result.

Lemma 4.2.9 (Spectral flow — see Ref. [205], Proposition 2.4)**.** *Let $\{H(x)\}_{x \in \mathcal{X}}$ be a uniformly gapped family of Hamiltonians with continuous first derivative with respect to x on the interval $\mathcal{X}$. Let $|\psi_0(x)\rangle$ denote the ground state of $H(x)$, and let $\gamma > 0$ be the uniform gap. Then the derivative of $|\psi_0(x)\rangle$ with respect to x is given by*

$$\frac{\mathrm{d}}{\mathrm{d}x} |\psi_0(x)\rangle = i \Psi_{H(x)}(\partial_x H(x)) |\psi_0(x)\rangle \ , \tag{4.34}$$

where $\Psi_H(\cdot)$ is the spectral flow operator defined by

$$\Psi_H(X) := \int_{-\infty}^{\infty} \mathrm{d}t w_\gamma(t) \int_0^t \mathrm{d}u e^{iuH} X e^{-iuH} \ , \tag{4.35}$$

and $w_\gamma(t)$ is any real-valued L_1 function satisfying $\int_{-\infty}^{\infty} \mathrm{d}t w_\gamma(t) = 1$ and whose Fourier transform $\hat{w}_\gamma$ is supported in the interval $[-\gamma, \gamma]$.

Note that both the quantum belief propagation and spectral flow operators $\Psi_H(X)$ correspond to a "smearing" of the operator X, under local dynamics controlled by H, and hence their supports are approximately localised around the support of X (we formalise this intuition with Lemma 4.2.15). Although we can intuitively view the spectral flow as the extension of quantum belief propagation to the $\beta \to \infty$ case, note that this is not how it is obtained (Eq. (4.28) is unbounded in this limit); the spectral gap is crucial.

In [205], an explicit family of such functions w_γ is given which decay quickly away from $t = 0$. Choosing this family, we can assume without loss of generality that $w_\gamma(t)$ is even, non-negative, and that when $|t| \geq e^{-1/\sqrt{2}}\gamma^{-1}$, we have the bounds

$$0 \leq w_\gamma(t) \leq 2(e\gamma)^2|t| \exp\left(-\frac{2\gamma|t|}{7\log^2(\gamma|t|)}\right) . \tag{4.36}$$

Moreover, for an observable O with ground state expectation value $f_{\mathsf{ground}}(x) := \langle\psi_0(x)|O|\psi_0(x)\rangle$, we can see from Eq. (4.34) that the derivative of $f_{\mathsf{ground}}(x)$ is given by

$$\frac{\mathrm{d}}{\mathrm{d}x} f_{\mathsf{ground}}(x) = \langle\psi_0(x)|i[O, \Psi_{H(x)}(\partial_x H(x))]|\psi_0(x)\rangle = \langle\psi_0(x)|i[\Psi_{H(x)}(O), \partial_x H(x)]|\psi_0(x)\rangle . \tag{4.37}$$

The absolute value of this derivative can thus be bounded by the operator norm of the commutator

$$\left[\Psi_{H(x)}(O), \partial_x H(x)\right] . \tag{4.38}$$

In Section 4.2.3, we will use Lieb-Robinson bounds to show that $\Psi_{H(x)}(O)$ is approximately localised around the support of O, and hence that Eq. (4.37) only depends on the variation of $H(x)$ supported near to O. For this purpose, it will be useful to prove the following alternative representation of the spectral flow operator:

Lemma 4.2.10 (Alternative representation of spectral flow — see Lemma 4.6.2)**.** *The spectral flow operator $\Psi_H(O)$ as defined in Lemma 4.2.9 can be written as*

$$\Psi_H(X) = \int_{-\infty}^{\infty} \tilde{w}_\gamma(t) e^{itH} X e^{-itH} , \tag{4.39}$$

where $\tilde{w}_\gamma : \mathbb{R} \to \mathbb{R}$ is an odd L_1 function with $|\tilde{w}_\gamma(t)| \leq 1/2$ for all t. Moreover, for $|t| \geq e^3\gamma^{-1}$,

$$|\tilde{w}_\gamma(t)| \leq W_1 \left(\frac{\gamma t}{\log^2(\gamma t)}\right)^2 \exp\left(-\frac{2\gamma t}{7\log^2(\gamma t)}\right) , \tag{4.40}$$

$$\int_t^\infty \mathrm{d}s \tilde{w}_\gamma(s) \leq W_2\gamma^{-1} \left(\frac{\gamma t}{\log^2(\gamma t)}\right)^3 \exp\left(-\frac{2\gamma t}{7\log^2(\gamma t)}\right) , \tag{4.41}$$

for some constants $W_1, W_2 > 0$. Moreover,

$$\int_{-\infty}^{\infty} \mathrm{d}t |\tilde{w}_\gamma(t)| \leq W_3\gamma^{-1} , \tag{4.42}$$

for some constant $W_3 > 0$.

4.2.3 Spread of correlations

Lieb-Robinson bounds

In order to analyse the locality properties of the quantum belief propagation and spectral flow operators introduced above, we require some general results about the localisation of operators under short-time dynamics [54]. We will essentially follow the formalism of Ref. [199], in which the authors prove very general results (which are also applicable to infinite lattices, though this case is not necessary for our purposes). See also Ref. [213], in which the authors use similar techniques with generalisation to dissipative dynamics.

Before we can state the Lieb-Robinson bounds, we adopt the terminology of Ref. [199] and define an F-function as below.

Definition 4.2.11 (F-function [199])**.** A non-increasing function $F : [0, \infty) \to [0, \infty)$ is called an F-function on (Γ, dist) with parameters $(\|F\|, C_F)$, where we define

$$\|F\| := \sup_{i \in \Gamma} \sum_{j \in \Gamma} F(\mathrm{dist}(i, j)) \ , \tag{4.43}$$

$$C_F := \sup_{i,j \in \Gamma} \sum_{k \in \Gamma} \frac{F(\mathrm{dist}(i, k))F(\mathrm{dist}(k, j))}{F(\mathrm{dist}(i, j))} \ . \tag{4.44}$$

When $\|F\|$ and C_F are both constants independent of the size of the lattice n, we will simply refer to F as an F-function. If $C_F = 1$, we say that F is normalised.

For our results in Sections 4.4-4.5, we will assume that all F-functions are normalised (and hence ignore any factors of C_F). This only affects our conclusions up to constant factors, since any F-function F can be be normalised via $F \mapsto C_F^{-1} F$.

The constants $\|F\|$ and C_F are both necessary to ensure that a Hamiltonian with bounded $\|\cdot\|_F$-norm is sufficiently localised to satisfy Lieb-Robinson bounds. Informally, $\|F\|$ gives a bound on the total strength of interactions acting on any given $i \in \Gamma$ for a Hamiltonian $H = \sum_{A \subseteq \Gamma} H_A$, as

$$\sum_{\substack{A \subseteq \Gamma \\ i \in A}} \|H_A\| \leq \sum_{j \in \Gamma} \sum_{\substack{A \subseteq \Gamma \\ i,j \in A}} \|H_A\| \tag{4.45}$$

$$\leq \sum_{j \in \Gamma} F(\mathrm{dist}(i, j)) \|H\|_F \tag{4.46}$$

$$\leq \|H\|_F \|F\| \ . \tag{4.47}$$

Meanwhile, the constant C_F will later be important to ensure that higher-order functions of local Hamiltonians remain local (for a concrete statement, see Lemma 4.6.7).

The phrase "constants independent of the size of the lattice" is not entirely well defined, as we consider a finite system Γ. This can be understood by viewing Γ as a member of a family of systems of varying size n, or as a subset of size n of the D-dimensional square lattice $\Gamma \subseteq \mathbb{Z}^D$. In either case, we then interpret these conditions as meaning that $\|F\|$ and C_F are bounded as n is taken infinitely large.

It is not immediately clear how to construct functions satisfying Definition 4.2.11, but there are two important examples for our purposes, as noted in [199]. Firstly, for the D-dimensional systems that we consider, the function $F(r) = (1+r)^{-(D+\epsilon)}$ defines an F-function for any $\epsilon > 0$, where $C_F \leq 2^{D+\epsilon}\|F\|$. Moreover, given any F-function F and a non-decreasing and subadditive (i.e. satisfying $g(r+s) \leq g(r) + g(s)$) function $g : [0,\infty) \to [0,\infty)$, the weighted function

$$F_g(r) := F(r)e^{-g(r)} \tag{4.48}$$

also defines an F-function. Choosing $g(r) = ar$ for constant a, this allows us to describe exponentially decaying interactions. We formalise this in the below definition:

Definition 4.2.12 (Exponentially decaying interactions)**.** A Hamiltonian H on Γ in D dimensions is said to have exponentially decaying interactions with decay rate $a > 0$ if $\|H\|_{F_g}$ is bounded by a constant, where F_g is the following F-function:

$$F_g(r) := (1+r)^{-(D+1)}e^{-ar} \ . \tag{4.49}$$

We can generalise the bound in Eq. (4.47) to bound the error incurred in operator norm when the Hamiltonian H is restricted to a sublattice $\Gamma' \subseteq \Gamma$, as below:

Lemma 4.2.13 (Hamiltonian restriction)**.** *Let H be a Hamiltonian on Γ with bounded $\|\cdot\|_F$-norm for an F-function F. For any subset $\Gamma' \subseteq \Gamma$, we define the restriction of H to Γ' by*

$$H|_{\Gamma'} := \sum_{A \subseteq \Gamma'} H_A \ . \tag{4.50}$$

Then the difference between H and $H|_{\Gamma'}$ is bounded by

$$\|H - H|_{\Gamma'}\| \leq |\Gamma \setminus \Gamma'|\|H\|_F\|F\| \ . \tag{4.51}$$

Proof of Lemma 4.2.13. We have

$$H - H|_{\Gamma'} = \sum_{\substack{A \subseteq \Gamma \\ A \cap (\Gamma \setminus \Gamma') \neq \emptyset}} H_A \ , \tag{4.52}$$

and hence

$$\|H - H|_{\Gamma'}\| \leq \sum_{x \in \Gamma \setminus \Gamma'} \sum_{\substack{A \subseteq \Gamma \\ x \in A}} \|H_A\| \tag{4.53}$$

$$\leq |\Gamma \setminus \Gamma'|\|H\|_F\|F\| \ . \tag{4.54}$$

where in the second line we used Eq. (4.47). ∎

With these definitions in hand, we are ready to state the Lieb-Robinson bounds we will use. We consider a time-dependent Hamiltonian $H(t) = \sum_{A \subseteq \Gamma} H_A(t)$, an F-function $F(r)$, and a function $g(r)$ as above. We use the shorthand $\|H\|_{F_g} := \sup_t \|H(t)\|_{F_g}$ (and assume that this is constant). Let O_A and O_B be two observables with disjoint supports $A, B \subseteq \Gamma$ separated by distance $\mathrm{dist}(A, B) = r$, and denote by $O_A(t)$ the time-evolved observable for $t \geq 0$, satisfying

$$O_A(0) = O_A \ , \quad \frac{\mathrm{d}}{\mathrm{d}t} O_A(t) = i[H(t), O_A(t)] \ . \tag{4.55}$$

Lemma 4.2.14 (Lieb-Robinson bounds — see Ref. [199], Theorem 3.1)**.** *There exist constants (depending on C_{F_g} and $\|F_g\|$) $c, \nu > 0$ such that, for all $t > 0$,*

$$\|[O_A(t), O_B]\| \leq c\|O_A\|\|O_B\| \min\{|A|, |B|\} \left(e^{\nu \|H\|_{F_g} t} - 1\right) e^{-g(r)} \ , \tag{4.56}$$

where $r = \mathrm{dist}(A, B)$.

Local truncations

We will now use the Lieb-Robinson bounds of the previous section to show that, for a local observable O and quasi-local Hamiltonian H, the operators $\Phi_H(O)$ and $\Psi_H(O)$ are localised around the support of O. For a set of sites $A \subseteq \Gamma$, we write $B_r(A)$ to denote the ball of radius r around A, that is

$$B_r(A) = \{i \in \Gamma : \mathrm{dist}(i, A) \leq r\} \ . \tag{4.57}$$

The below result, Lemma 4.2.15, establishes that whenever O_A is supported in $A \subseteq \Gamma$, $\Phi_H(O_A)$ and $\Psi_H(O_A)$ have good approximations which only act on $B_r(A)$. See also Ref. [193], Lemma III.2 for essentially the same result in the case of quantum belief propagation (a similar statement also appears in Ref. [211], Corollary 3), and Ref. [199] Section 6.5 for similar results establishing quasi-locality of the spectral flow. We defer the proof of this result to Appendix 4.6.2.

Lemma 4.2.15 (Truncation of the quantum belief propagation and spectral flow operators — see Lemma 4.6.4)**.** *Let H be a Hamiltonian with bounded $\|\cdot\|_{F_g}$-norm, and let O_A be an observable supported on $A \subseteq \Gamma$. Let $\Phi_H(O_A)$ be the quantum belief propagation operator defined by Eq. (4.29), and $\Psi_H(O_A)$ be the spectral flow operator defined by Eq. (4.35). Then for every $r \geq 0$ there exist operators $\Phi_H^{[r]}(O_A)$ and $\Psi_H^{[r]}(O_A)$ which act as the identity outside of $B_r(A)$, and constants $a_1, b_1 > 0$ (depending on $c, \beta, \nu, \|H\|_{F_g}$) and $a_2, b_2 > 0$ (depending on $c, \gamma, \nu, \|H\|_{F_g}$), such that*

$$\|\Phi_H(O_A) - \Phi_H^{[r]}(O_A)\| \leq a_1 |A| \|O_A\| e^{-b_1 g(r)} \ , \tag{4.58}$$

$$\|\Psi_H(O_A) - \Psi_H^{[r]}(O_A)\| \leq a_2 |A| \|O_A\| e^{-b_2 g(r)/\log^2 g(r)} \ , \tag{4.59}$$

where $\omega(x)$ is defined by

$$\omega(x) := \left(\frac{x}{\log^2(x)}\right)^3 \exp\left(-\frac{2x}{7\log^2 x}\right) \ . \tag{4.60}$$

Generalised approximate local indistinguishability

This ability to truncate the action of the quantum belief propagation and spectral flow operators allows us to formalise the notion that local observables are not strongly affected by a distantly changing Hamiltonian in thermal and ground states away from criticality. This phenomenon is known as generalised approximate local indistinguishability (GALI), and was shown in Ref. [193] to hold for Gibbs states with exponentially decaying correlations and gapped ground states. The results below are essentially an adaptation of their results to our framework.

Lemma 4.2.16 (GALI for Gibbs states — [193] Proposition V.4 paraphrased — see Lemma 4.6.5)**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ be a family of Hamiltonians with continuous first derivative with respect to x on the interval $\mathcal{X}$, such that $\|H\|_{F_g}$ and $\|\partial_x H\|_{F_g}$ are both bounded. Let O_A be an observable supported on $A \subseteq \Gamma$, and assume that $\partial_s H(s)$ contains no terms with support in $B_{r_0}(A)$ for some $r_0 \geq 0$.*

Let $\rho_\beta(x)$ be the associated family of Gibbs states at temperature $\beta = \mathcal{O}(1)$, and assume that these satisfy a uniform exponential decay of correlations with parameters $K, \xi > 0$. Then, for all x, it holds that

$$\left|\mathrm{Cov}_{\rho_\beta(x)}(\Phi_{H(x)}(O_A), \partial_x H(x))\right| \leq c_1 \|O_A\||A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)}\right) , \quad (4.61)$$

for positive constants $b_1, c_1 > 0$. Hence, defining $f_\beta(x) := \mathrm{tr}[O_A \rho_\beta(x)]$, we have for all $x_0, x_1 \in \mathcal{X}$ that

$$|f_\beta(x_1) - f_\beta(x_0)| \leq \beta |x_1 - x_0| c_1 \|O_A\||A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)}\right) . \quad (4.62)$$

Notice that in the case of exponentially decaying correlations (that is, when $g(r) = \Omega(r)$), the right-hand side of Eq. (4.62) decays exponentially with r_0. This will allow us to estimate $f_\beta(x)$ to accuracy $\epsilon > 0$ whilst ignoring the variation of $H(x)$ outside a ball of radius $r_0 \sim \log \epsilon^{-1}$. We state the corresponding result for gapped ground states below.

Lemma 4.2.17 (GALI for ground states — [193] Proposition V.5 paraphrased (see also [194]) — see Lemma 4.6.6)**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ and O_A be as in Lemma 4.2.16, and assume that $\{H(x)\}_{x\in\mathcal{X}}$ has a uniform gap $\gamma > 0$ above its ground state $|\psi_0(x)\rangle$. Then*

$$\left\|[\Psi_{H(x)}(O_A), \partial_x H(x)]\right\| \leq c_2 \|O_A\||A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} , \quad (4.63)$$

for positive constants $b_2, c_2 > 0$. Hence, defining $f_{\mathsf{ground}}(x) := \langle\psi_0(x)|O_A|\psi_0(x)\rangle$, we have for all $x_0, x_1 \in \mathcal{X}$ that

$$|f_{\mathsf{ground}}(x_1) - f_{\mathsf{ground}}(x_0)| \leq |x_1 - x_0| c_2 \|O_A\||A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} . \quad (4.64)$$

In particular, in the case of exponential decaying interactions $g(r) = \Omega(r)$, the right-hand side of Eq. (4.64) decays as $\sim \exp(-\Omega(r/\log^2 r))$.

4.3 Extrapolation within phases of matter

4.3.1 Assumptions and main statement

In this section, we establish the first set of results for this work: namely, that local properties of Gibbs states and ground states have good analytic approximations (and hence can be extrapolated) along paths of non-critical Hamiltonians. This will establish some basic results and proof techniques which we will ultimately combine with the perturbation theory machinery in Section 4.4 to prove our main results about extrapolating simulator Hamiltonians in Section 4.5. The main conclusions of this section are stated below as Theorem 4.3.2. The proof of the result is divided into its two parts: Theorems 4.3.9 and 4.3.14 deal with the Gibbs state and ground state results respectively. The results are then proved separately in Sections 4.3.2 and 4.3.3. A generalisation of Theorem 4.3.2 to the case of extensive observables is given in the appendix; see Theorem 4.6.17.

Firstly, we make precise the non-criticality assumptions required for our main results, collected below as Assumption 4.3.1 for convenience. In addition to geometric locality and analytic dependence on the parameter x, we assume that the family of Hamiltonians has a uniform exponential decay of correlations (respectively a spectral gap, for the ground state case) which is robust even when the interactions contained within some region $\Gamma' \subseteq \Gamma$ are parametrised differently to those outside of Γ'. Ultimately, this will be necessary for us to argue that we can continuously "turn off" the interactions far from the local observable of interest, using the GALI properties from Section 4.2.3.

Assumption 4.3.1. *Let $H(x) = \sum_{A \subseteq \Gamma} H_A(x)$ be a family of Hamiltonians parametrised by $x \in \mathcal{X}$. Assume that each $H_A(x)$ is an analytic function of x, and that both $H(x)$ and $\partial_x H(x)$ have exponentially decaying interactions as in Definition 4.2.12 (that is, $\|H(x)\|_{F_g}, \|\partial_x H(x)\|_{F_g} = \mathcal{O}(1)$ for an F-function F and for $g(r) = ar$ where $a > 0$) whenever $x \in \{z \in \mathbb{C} : |z| \leq x_*\}$, for some $x_* > 0$. For every $\Gamma' \subseteq \Gamma$, define the family of Hamiltonians $\{H(x, y; \Gamma')\}_{(x,y) \in \mathcal{X} \times \mathcal{X}}$ by*

$$H(x, y; \Gamma') := \sum_{A \subseteq \Gamma'} H_A(x) + \sum_{\substack{A \subseteq \Gamma \\ A \nsubseteq \Gamma'}} H_A(y) \ . \tag{4.65}$$

Then we further assume that one of the following holds, depending on the extrapolation task:

(I) (For Gibbs states)*: There exist constants $K, \xi > 0$ such that, for all $\Gamma' \subseteq \Gamma$, the family $\{\rho_\beta(x, y; \Gamma')\}_{(x,y) \in \mathcal{X} \times \mathcal{X}}$ satisfies a uniform exponential decay of correlations with parameters K, ξ (see Definition 4.2.7).*

(II) (For ground states)*: There exists a constant $\gamma > 0$ such that, for all $\Gamma' \subseteq \Gamma$, the family $\{H(x, y; \Gamma')\}_{(x,y) \in \mathcal{X} \times \mathcal{X}}$ is uniformly gapped above the unique ground state with gap γ (see Definition 4.2.5).*

We can now state the main result of the section. This is stated in terms of analytic approximations of local observables: for any chosen $\delta > 0$, we establish the existence of (δ, M, R)-analytic approximations, for $M, R > 0$ depending on δ.

Theorem 4.3.2 (See Theorems 4.3.9 and 4.3.14)**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ be a family of Hamiltonians on Γ in D dimensions satisfying the conditions of Assumption 4.3.1. Let O_A be an observable supported on $A \subseteq \Gamma$, where $|A|, \|O_A\| = \mathcal{O}(1)$. Then the following holds:*

(I) Let $\rho_\beta(x) \sim e^{-\beta H(x)}$ be the family of Gibbs states associated to $H(x)$ at inverse temperature $\beta > 0$. Then (with Assumption 4.3.1(I)*) the function*

$$f_\beta(x) := \mathrm{tr}[O_A \rho_\beta(x))] \ , \quad x \in [0, x_*] \ , \tag{4.66}$$

has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = 3\|O_A\| \ , \quad x_* \geq R = 1/\,\mathcal{O}(\log^D(\delta^{-1})) \ . \tag{4.67}$$

(II) Let $|\psi_0(x)\rangle$ be the unique ground state of $H(x)$. Then (with Assumption 4.3.1(II)*) the function*

$$f_{\mathsf{ground}}(x) := \langle \psi_0(x))|O_A|\psi_0(x)\rangle \ , \quad x \in [0, x_*] \ , \tag{4.68}$$

has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = 16\|O_A\| \ , \quad x_* \geq R = 1/\,\mathcal{O}(\log^{D+1}(\delta^{-1})) \ . \tag{4.69}$$

In particular, this theorem implies — by Corollary 4.2.4 — that the functions $f_\beta(x)$ and $f_{\mathsf{ground}}(x)$ can be extrapolated to $x = 0$ with accuracy $\epsilon > 0$ from their evaluations at values of x lower bounded by $1/\operatorname{poly}\log(\epsilon^{-1})$. We make this concrete with the following corollary:

Corollary 4.3.3 (Extrapolating local observables within phases of matter)**.** *The value of $f_\beta(0)$ (respectively $f_{\mathsf{ground}}(0)$) can be calculated up to any desired accuracy $\epsilon > 0$, given the values of $f_\beta(x_k)$ (respectively $f_{\mathsf{ground}}(x_k)$) at m Chebyshev sample points $\{x_k\}_{k=1}^m$, where each x_k is bounded above zero by $x_{\min} := \min_k x_k$, where:*

(I) For f_β,

$$m = \mathcal{O}(\log(\epsilon^{-1})) \ , \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{D+2}(\epsilon^{-1}))} \ . \tag{4.70}$$

(II) For f_{ground},

$$m = \mathcal{O}(\log(\epsilon^{-1})) \ , \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{D+3}(\epsilon^{-1}))} \ . \tag{4.71}$$

The result still holds using noisy estimates of $f_\beta(x_k)$ (respectively $f_{\mathsf{ground}}(x_k)$) each with additive error $\delta = \Theta(\epsilon/\log\log(\epsilon^{-1}))$.

Proof of Corollary 4.3.3. Corollary 4.2.4 ensures that we can extrapolate to within error $\epsilon > 0$, given a (δ, M, R)-analytic approximation and m Chebyshev samples, as long as

$$\epsilon = (\delta + 2^{-m} M)\, \mathcal{O}(\log m) \ . \tag{4.72}$$

For both f_β and f_{ground}, we can obtain the required approximations from Theorem 4.3.2 with $M = \mathcal{O}(1)$, and hence a choice of $m = \mathcal{O}(\log(\epsilon^{-1}))$ and $\delta = \Theta(\epsilon / \log\log(\epsilon^{-1}))$ suffices. This leads to $R = 1/\,\mathcal{O}(\log^D(\epsilon^{-1}))$ and $R = 1/\,\mathcal{O}(\log^{D+1}(\epsilon^{-1}))$ for f_β and f_{ground} respectively. We then have, from Corollary 4.2.4, that $x_{\min} \sim R/m^2$, from which the stated result follows. ∎

4.3.2 Gibbs states

Analyticity of perturbations on Gibbs states

In order to establish the Gibbs state case of Theorem 4.3.2, the proof will proceed in two parts. Firstly, we use the GALI property of Lemma 4.2.16 to show that the Hamiltonian perturbation can be approximately truncated to a spatially localised region. Next, we need to show that sufficiently small (in operator norm) perturbations to the Hamiltonian result in perturbations to the Gibbs state which can be extended analytically. To this end, we need to prove the following theorem:

Theorem 4.3.4. *Let $H \in \mathrm{Herm}(\mathcal{H})$ be a Hamiltonian, and let $V : \mathbb{C} \to \mathrm{Lin}(\mathcal{H})$ be an analytic family of perturbations. Let $\beta > 0$, and let $R > 0$ be such that $V(z)$ is bounded as*

$$\|V(z)\| \leq \beta^{-1} \log(3/2) \quad \textit{whenever} \quad |z| \leq R \ . \tag{4.73}$$

Fix an observable $O \in \mathrm{Herm}(\mathcal{H})$. Then the function

$$f_\beta(z) := \frac{\mathrm{tr}[O e^{-\beta(H+V(z))}]}{\mathrm{tr}[e^{-\beta(H+V(z))}]} \tag{4.74}$$

is analytic in the disc $|z| \leq R$, and moreover is bounded by

$$\sup_{|z| \leq R} |f_\beta(z)| \leq 3\|O\| \ . \tag{4.75}$$

(Proof on page 124.)

Lemma 4.3.5 is the key technical fact for this purpose; it is quite a general trace inequality which may be of independent interest:

Lemma 4.3.5. *Let $X \in \mathrm{Herm}(\mathcal{H})$ be a Hermitian matrix, and let $Y \in \mathrm{Lin}(\mathcal{H})$ be a general matrix. Let $\omega(X, Y) = |\,\mathrm{tr}[e^{X+Y}]|/|\,\mathrm{tr}[e^X]|$. Then*

$$2 - e^{\|Y\|} \leq \omega(X, Y) \leq e^{\|Y\|} \ . \tag{4.76}$$

We think of $X = -\beta H$ proportional to an unperturbed Hamiltonian, and $Y = -\beta V(z)$ being a family of perturbations depending on a complex parameter z — thus $\omega(X, Y)$ has the interpretation as the ratio of the corresponding perturbed partition function to an unperturbed partition function. The fact that $\omega(X, Y) > 0$ for some region of z is sufficient to deduce that $\mathrm{tr}[e^{-\beta(H+V(z))}]$ has no complex zeroes in this region, but is not sufficient to bound the Taylor series truncation error of the corresponding Gibbs state. For this, we need an explicit positive lower bound for $\omega(X, Y)$, ensuring that the perturbed partition function cannot get "too small" and hence that expectation values of the perturbed ground state are bounded for the region of complex z. We note that the lower bound in (4.76) is not tight (and is only non-trivial when $\|Y\| \leq \log 2$) but it is sufficient for our purposes.

Proof of Lemma 4.3.5. The upper bound is immediate from Lemma 4.3.8, since

$$\omega(X, Y) = \frac{|\,\mathrm{tr}[e^{X+Y}]|}{\mathrm{tr}[e^X]} \leq \frac{\|e^{X+Y}\|_1}{\mathrm{tr}[e^X]} \leq e^{\|Y\|} \,. \tag{4.77}$$

For the lower bound, we introduce a function $f(t)$ defined by

$$f(t) = \mathrm{Re}\left(\frac{\mathrm{tr}[e^{X+Yt}]}{\mathrm{tr}[e^X]}\right), \quad t \in [0, 1] \,. \tag{4.78}$$

Note that $f(0) = 1$ and $f(1) \leq \omega(X, Y)$. Differentiating, we have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(t) &= \frac{\mathrm{Re}\,\mathrm{tr}[Y e^{X+Yt}]}{\mathrm{tr}[e^X]} \\ &\geq -\frac{\|Y\| \|e^{X+Yt}\|_1}{\mathrm{tr}[e^X]} \\ &\geq -\|Y\| e^{\|Y\|t} \,, \end{aligned} \tag{4.79}$$

where in the second line we have used Hölder's inequality and in the third line we have applied Lemma 4.3.8. Hence

$$f(1) = f(0) - \int_0^1 \mathrm{d}t \frac{\mathrm{d}}{\mathrm{d}t} f(t) \geq 1 - \|Y\| \int_0^1 \mathrm{d}t e^{\|Y\|t} = 2 - e^{\|Y\|} \,. \tag{4.80}$$

■

We now show how Lemma 4.3.5 implies Theorem 4.3.4. We will need the following two lemmas on the singular values of matrices.

Lemma 4.3.6 ([214], Corollary 2.3)**.** *Let $X \in \mathrm{Lin}(\mathcal{H})$ be an arbitrary complex matrix. Let $\{s_j(X)\}_j$ be the singular values of X, and let $\{s_j(e^X)\}_j$ be the singular values of e^X. Then*

$$\sum_j s_j(e^X) \leq \sum_j e^{s_j(X)} \,. \tag{4.81}$$

Lemma 4.3.7 (Weyl's inequality for singular values)**.** *Let X and Y be arbitrary matrices, with singular values $0 \leq s_0(X) \leq s_1(X) \leq \dots$ and $0 \leq s_0(Y) \leq s_1(Y) \leq \dots$ respectively. Then*

$$\max_j |s_j(X) - s_j(Y)| \leq \|X - Y\| \ . \tag{4.82}$$

From here we can prove the following useful bound on the 1-norm of a matrix exponential.

Lemma 4.3.8. *Let $X \in \mathrm{Herm}(\mathcal{H})$ be Hermitian, and let $Y \in \mathrm{Lin}(\mathcal{H})$ be an arbitrary matrix. Then*

$$\|e^{X+Y}\|_1 \leq e^{\|Y\|} \operatorname{tr}[e^X] \ . \tag{4.83}$$

Proof of Lemma 4.3.8. Using the notation of Lemma 4.3.6, we have

$$\begin{aligned} \|e^{X+Y}\|_1 &= \sum_j s_j(e^{X+Y}) \\ &\leq \sum_j e^{s_j(X+Y)} \\ &\leq e^{\|Y\|} \sum_j e^{s_j(X)} \\ &= e^{\|Y\|} \operatorname{tr}[e^X] \ , \end{aligned} \tag{4.84}$$

where in the second line we used Lemma 4.3.6 and in the third line we used Lemma 4.3.7. ∎

Proof of Theorem 4.3.4. The bound (4.73), along with Lemma 4.3.5, guarantees that $\operatorname{tr}[e^{-\beta(H+V(z))}]$ has no zeroes in the region $|z| \leq R$, and so (using the analyticity of V) it is immediate that $f_\beta(z)$ is analytic in this disc.

For the upper bound, we can compute

$$|f_\beta(z)| = \left|\frac{\operatorname{tr}[e^{-\beta H}]}{\operatorname{tr}[e^{-\beta(H+V(z))}]}\right| \cdot \left|\frac{\operatorname{tr}[Oe^{-\beta(H+V(z))}]}{\operatorname{tr}[e^{-\beta H}]}\right| \leq (2 - e^{\|\beta V(z)\|})^{-1} \|O\| \frac{\|e^{-\beta(H+V(z))}\|_1}{\operatorname{tr}[e^{-\beta H}]} \ , \tag{4.85}$$

where we bounded the first term using Lemma 4.3.5 and the second term using Hölder's inequality. Using Lemma 4.3.8, we can conclude

$$|f_\beta(z)| \leq (2 - e^{\|\beta V(z)\|})^{-1} \|O\| e^{\|\beta V(z)\|} \leq 3\|O\| \ , \tag{4.86}$$

where in the second inequality we used that $e^x/(2 - e^x) \leq 3$ for $0 \leq x \leq \log(3/2)$. ∎

Analytic approximation of observables on non-critical Gibbs states

We are now ready to prove the Gibbs state part of Theorem 4.3.2. We restate the result as the following theorem:

Theorem 4.3.9 (Extrapolation of local properties of Gibbs states — restatement of Theorem 4.3.2(I))**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ be a family of Hamiltonians satisfying Assumption 4.3.1*(I)*. Then, for any $\delta > 0$, the function $f_\beta(x)$ defined in Eq.* (4.66) *has a (δ, M, R)-analytic approximation, where*

$$M = 3\|O_A\| \ , \quad x_* \geq R = 1/\,\mathcal{O}(\log^D(\delta^{-1})) \ . \tag{4.87}$$

Proof of Theorem 4.3.9. Let $r_0 > 0$ (which will later be set to $r_0 \sim \log \delta^{-1}$), and write $f_\beta(x, y; B_{r_0}(A)^c)$ for the function

$$f_\beta(x, y; B_{r_0}(A)^c) := \mathrm{tr}[O_A \rho_\beta(x, y; B_{r_0}(A)^c)] \ , \tag{4.88}$$

where $\rho_\beta(x, y; B_{r_0}(A)^c)$ is defined, as in Assumption 4.3.1(I), as the Gibbs state corresponding to the Hamiltonian

$$H(x, y; B_{r_0}(A)^c) := \sum_{A' \subseteq B_{r_0}(A)^c} H_{A'}(x) + \sum_{\substack{A' \subseteq \Gamma \\ A' \not\subseteq B_{r_0}(A)^c}} H_{A'}(y) \ . \tag{4.89}$$

That is, the Hamiltonian where all interactions whose support intersects $B_{r_0}(A)$ are parametrised by y, whilst all others are parametrised by x. By the GALI property for Gibbs states (Lemma 4.2.16), $f_\beta(x, y; B_{r_0}(A)^c)$ only depends very slightly on its first argument, in the sense that we can bound:

$$|f_\beta(x, x; B_{r_0}(A)^c) - f_\beta(0, x; B_{r_0}(A)^c)| \leq \beta|x|\|O_A\||A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 a r/2} \right) \tag{4.90}$$

$$\leq |x| a' e^{-b' r_0} \ , \tag{4.91}$$

for some constants $b_1, c_1 > 0$, and where on the second line we have observed that the summand is exponentially decaying in r_0 and absorbed all the constants in the expression into the constants $a', b' > 0$.

We write $\tilde{f}_\beta(x) := f_\beta(0, x; B_{r_0}(A)^c)$. For any complex $|x| \leq x_*$, the Hamiltonian

$H(0,x;B_{r_0}(A)^c)$ can be viewed as a perturbation from $H(0)$ of magnitude

$$\begin{aligned}
\|H(0,x;B_{r_0}(A)^c) - H(0)\| &\leq \sum_{\substack{A'\subseteq\Gamma \\ A'\cap B_{r_0}(A)\neq\emptyset}} \|H_{A'}(x) - H_{A'}(0)\| && (4.92)\\
&\leq \sum_{i\in B_{r_0}(A)} \sum_{\substack{A'\subseteq\Gamma \\ i\in A'}} \|H_{A'}(x) - H_{A'}(0)\| && (4.93)\\
&\leq \sum_{i\in B_{r_0}(A)} \sum_{\substack{A'\subseteq\Gamma \\ i\in A'}} |x| \sup_{|x'|\leq R} \left\|\partial_{x'} H_{A'}(x')\right\| && (4.94)\\
&\leq |B_{r_0}(A)||x| \sup_{|x'|\leq R} \|\partial_{x'} H(x')\|_{F_g} \|F_g\| \ . && (4.95)
\end{aligned}$$

In the first and second lines we have used the triangle inequality, in the third line we used the mean value theorem, and in the fourth line we have used the definition of the $\|\cdot\|_{F_g}$-norm. Now, since $|B_{r_0}(A)| \leq |A| k_D r_0^D$, and treating $|A|, k_D, D, \|F_g\|$, and $\sup_{|s'|\leq R} \|\partial_{s'} H(s')\|_{F_g}$ as constants, we have

$$\|H(0,x;B_{r_0}(A)^c) - H(0)\| \leq c'|x| r_0^D \ , \tag{4.96}$$

for some constant $c' > 0$. Thus, using Theorem 4.3.4 with $V(x) := H(0,x;B_{r_0}(A)^c) - H(0)$, we deduce that $\tilde{f}_\beta(x)$ is analytic and has $\tilde{f}_\beta(x) \leq 3\|O_A\|$ in the complex disc defined by

$$c'|x| r_0^D \leq \beta^{-1} \log(3/2) \Rightarrow |x| \leq R = \frac{\log(3/2)}{\beta b' r_0^D} = \mathcal{O}(r_0^{-D}) \ . \tag{4.97}$$

In other words, $\tilde{f}_\beta(x)$ is a (δ, M, R)-analytic approximation for $f_\beta(x)$, where

$$\delta = \exp(-\Theta(r_0)) \ , \quad M = 3\|O_A\| \ , \quad R = \Theta(r_0^{-D}) \ . \tag{4.98}$$

Hence, for any chosen $\delta > 0$ we may fix $r_0 = \mathcal{O}(\log(\delta^{-1}))$ to obtain the required result. ∎

4.3.3 Gapped ground states

Approximating the spectral flow

The proof idea for the gapped ground state part of Theorem 4.3.2 is similar to the Gibbs state case: we use GALI to assume that the Hamiltonian perturbations only occur locally (and therefore have bounded norm), and then prove that the effect of such small perturbations on observables can be analytically extended on the complex plane. The second part of this argument quickly runs into a fundamental problem: the spectral flow operator, defined through Lemma 4.2.10 as

$$\Psi_H(X) = \int_{-\infty}^{\infty} \tilde{w}_\gamma(t) e^{itH} X e^{-itH} \ , \tag{4.99}$$

is ill-defined whenever H is not Hermitian, as the integral will not necessarily converge. As a result, even very small complex Hamiltonian perturbations may lead to singular behaviour. Below, we prove a few preliminary results to circumvent this issue. Namely, in Lemma 4.3.10, we show that the integral in Eq. (4.99) can be truncated to finite range whilst incurring some mild errors. The following results, Lemmas 4.3.11 and 4.3.12, ensure that this approximate version of the spectral flow operator can be used to construct the desired analytic approximation to the function of interest $f_{\mathsf{ground}}(s)$.

Lemma 4.3.10 (Integral truncation for spectral flow)**.** *Let $H \in \mathrm{Herm}(\mathcal{H})$ be a Hamiltonian, and let $V : \mathbb{R} \to \mathrm{Herm}(\mathcal{H})$ be an family of Hamiltonian perturbations which extends to an analytic function $V : \mathbb{C} \to \mathrm{Lin}(\mathcal{H})$. For any analytic $X : \mathbb{C} \to \mathrm{Lin}(\mathcal{H})$, let*

$$\Psi_{H+V(x)}(X(x)) = \int_{-\infty}^{\infty} \mathrm{d}t \tilde{w}_\gamma(t) e^{it(H+V(x))} X(x) e^{-it(H+V(x))} \ , \quad x \in \mathbb{R} \ , \tag{4.100}$$

be the spectral flow operator as in Lemma 4.2.10. For any $T \geq 0$, define the truncated spectral flow operator by

$$\tilde{\Psi}_{H+V(z)}(X(z)) := \int_{-T}^{T} \mathrm{d}t \tilde{w}_\gamma(t) e^{it(H+V(z))} X(z) e^{-it(H+V(z))} \ , \quad z \in \mathbb{C} \ . \tag{4.101}$$

Then $\tilde{\Psi}_{H+V(z)}(X(z))$ is an analytic function of z, bounded by

$$\|\tilde{\Psi}_{H+V(z)}(X(z))\| \leq 2T\|X(z)\| \exp(2T\|V(z)\|) \ , \quad z \in \mathbb{C} \ , \tag{4.102}$$

and approximates $\Psi_{H+V(x)}(X(x))$ for real x up to error

$$\|\tilde{\Psi}_{H+V(x)}(X(x)) - \Psi_{H+V(x)}(X(x))\| \leq 2\|X(x)\| W_2 \gamma^{-1} \left(\frac{\gamma T}{\log^2(\gamma T)} \right)^3 \exp\left(-\frac{2\gamma T}{7 \log^2(\gamma T)} \right) \ , \quad x \in \mathbb{R} \ , \tag{4.103}$$

where $W_2 > 0$ is the constant from Lemma 4.2.10.

Proof of Lemma 4.3.10. The analyticity of $\tilde{\Psi}_{H+V(z)}(X(z))$ follows from the analyticity of the exponential function, and the standard fact that a weighted integral of analytic functions is itself analytic (see, e.g., Ref. [215] Theorem 5.4). Moreover, the bound follows from the triangle inequality

$$\left\| \tilde{\Psi}_{H+V(z)}(X(z)) \right\| \leq \int_{-T}^{T} \mathrm{d}t \left\| \tilde{w}_\gamma(t) e^{it(H+V(z))} X(z) e^{-it(H+V(z))} \right\| \tag{4.104}$$

$$\leq 2T\|X(z)\| \exp\left(2T\|V(z)\|\right) \ , \tag{4.105}$$

where in the second line we used that $|\tilde{w}_\gamma(t)| \leq 1$ (from Lemma 4.2.10), and the fact that $H \in \mathrm{Herm}(\mathcal{H})$. Finally, we can use the triangle inequality again, along with the

fact that $\tilde{w}_\gamma$ is an odd function, to bound

$$\|\tilde{\Psi}_{H+V(x)}(X(x)) - \Psi_{H+V(x)}(X(x))\| \leq 2\|X(x)\| \int_T^\infty \mathrm{d}t |\tilde{w}_\gamma(t)| \tag{4.106}$$

$$\leq 2\|X(x)\| W_2 \gamma^{-1} \left(\frac{\gamma T}{\log^2(\gamma T)}\right)^3 \exp\left(-\frac{2\gamma T}{7\log^2(\gamma T)}\right) , \tag{4.107}$$

as required, where in the second line we used the upper bound from Lemma 4.2.10. ∎

The next two results concern time-ordered exponentials and pure state evolution. Given a time-dependent Hamiltonian $H : \mathbb{R} \to \mathrm{Herm}(\mathcal{H})$, we define the time-ordered exponential by

$$U(x) = \exp_{\mathcal{T}}\left(i\int_0^x H(t)\,\mathrm{d}t\right) := \sum_{q\geq 0} i^q \int_{0\leq t_1\leq\cdots\leq t_q\leq x} \mathrm{d}t_1 \ldots \mathrm{d}t_q H(t_q)\ldots H(t_1) . \tag{4.108}$$

Then, given a family of states $|\psi(x)\rangle$ obeying

$$\frac{\mathrm{d}}{\mathrm{d}x}|\psi(x)\rangle = iH(x)|\psi(x)\rangle , \quad x \in \mathbb{R} , \tag{4.109}$$

the unique solution for $|\psi(x)\rangle$ given a starting state $|\psi(0)\rangle$ is

$$|\psi(x)\rangle = U(x)|\psi(0)\rangle . \tag{4.110}$$

The next result establishes that $U(x)$ (and hence $|\psi(x)\rangle$) extends to an analytic function, provided that $H(x)$ does.

Lemma 4.3.11 (Analyticity of time-ordered exponentials)**.** *Let $H : \mathbb{C} \to \mathrm{Lin}(\mathcal{H})$ be an analytic matrix-valued function in the disc $D_R := \{z \in \mathbb{C} : |z| \leq R\}$, which is bounded as $\sup_{|z|\leq R}\|H(z)\| = M$. Then $U(s)$ (as defined in Eq. (4.108)) extends to a function $U : \mathbb{C} \to \mathrm{Lin}(\mathcal{H})$ which is analytic for $|z| \leq R$, and which satisfies*

$$\sup_{|z|\leq R}\|U(z)\| \leq e^{MR} . \tag{4.111}$$

Proof of Lemma 4.3.11. We aim to define, for $z \in D_R$,

$$U(z) := \sum_{q\geq 0} i^q \int_0^z \mathrm{d}t_1 \int_{t_1}^z \mathrm{d}t_2 \cdots \int_{t_{q-1}}^z \mathrm{d}t_q H(t_q)\ldots H(t_1) , \tag{4.112}$$

though we must check that this is well-defined, in the sense that the integrals are independent of the path taken through $\mathbb{C}$ (or at least D_R). To this end, we prove by induction that each term of the form

$$\int_0^z \mathrm{d}t_1 \int_{t_1}^z \mathrm{d}t_2 \cdots \int_{t_{q-1}}^z \mathrm{d}t_q H(t_q)\ldots H(t_1) \tag{4.113}$$

is an analytic function of z, and independent of the path that t_1 takes. Indeed, this is trivially true for $q = 1$ by the analyticity of H. Moreover, we can write this as

$$\int_0^z \mathrm{d}t_1 \int_{t_1}^z \mathrm{d}t_2 \cdots \int_{t_{q-1}}^z \mathrm{d}t_q H(t_q) \dots H(t_1) = \int_0^z \mathrm{d}t_1 \Bigg(\int_0^z \mathrm{d}t_2 \cdots \int_{t_{q-1}}^z \mathrm{d}t_q H(t_q) \dots H(t_2) \tag{4.114}$$

$$- \int_0^{t_1} \mathrm{d}t_2 \cdots \int_{t_{q-1}}^z \mathrm{d}t_q H(t_q) \dots H(t_2) \Bigg) H(t_1) \ . \tag{4.115}$$

By induction, the bracketed term is an analytic function of both z and t_1, and — using the analyticity of H — so is the entire integrand. Since $U(z)$ is a sum of such terms, whose magnitude decays quickly (as $\sim \|zH\|^q/q!$), this proves that $U(z)$ is both well-defined and analytic (see Ref. [215], Theorem 5.2).

Using the independence of paths in the integral, we can without loss of generality choose the straight-line paths

$$U(|z|e^{i\theta}) = \sum_{q \geq 0} i^q e^{iq\theta} \int_0^{|z|} \mathrm{d}s_1 \int_{s_1}^{|z|} \mathrm{d}s_2 \cdots \int_{s_{q-1}}^{|z|} \mathrm{d}s_q H(s_q e^{i\theta}) \dots H(s_1 e^{i\theta}) \ . \tag{4.116}$$

Thus, using that the total volume of the integral through $\mathbb{R}^q$ is $|z|^q/q!$, we can bound

$$\|U(z)\| = \sum_{q \geq 0} \frac{|z|^q}{q!} \max_{w \leq |z|} \|H(w)\|^q \leq e^{M|z|} \ , \tag{4.117}$$

from which the result follows. ∎

The next result bounds the difference between the trajectories of two states evolving under slightly different Hamiltonians (for our purposes, this will be relevant to bound the evolution error between an ideal spectral flow operator, and its truncated approximation).

Lemma 4.3.12 (Perturbed pure state evolution)**.** *Let $H, V : \mathbb{R} \to \mathrm{Herm}(\mathcal{H})$ be two continuous families of Hamiltonians. Fix an initial state $|\psi_0\rangle \in \mathcal{H}$, and define the trajectories $|\psi(x)\rangle$ and $|\tilde{\psi}(x)\rangle$ by*

$$|\psi(0)\rangle = |\tilde{\psi}(0)\rangle = |\psi_0\rangle \ , \tag{4.118}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}|\psi(x)\rangle = iH(x)|\psi(x)\rangle \ , \tag{4.119}$$

$$\frac{\mathrm{d}}{\mathrm{d}x}|\tilde{\psi}(x)\rangle = i(H(x) + V(x))|\tilde{\psi}(x)\rangle \ . \tag{4.120}$$

Then, for all $x \in \mathbb{R}$,

$$\Big\| |\psi(x)\rangle - |\tilde{\psi}(x)\rangle \Big\| \leq \int_0^x \mathrm{d}t \|V(t)\| \ . \tag{4.121}$$

Proof of Lemma 4.3.12. We define the vector $|\epsilon(x)\rangle := |\psi(x)\rangle - |\tilde{\psi}(x)\rangle$. We can compute

$$\frac{\mathrm{d}}{\mathrm{d}x}|\epsilon(x)\rangle = iH(x)|\psi(x)\rangle - i(H(x)+V(x))|\tilde{\psi}(x)\rangle \tag{4.122}$$

$$= i(H(x)+V(x))|\epsilon(x)\rangle - iV(x)|\psi(x)\rangle \ . \tag{4.123}$$

One may verify that the unique solution of this differential equation with $|\epsilon(0)\rangle = 0$ is given by

$$|\epsilon(x)\rangle = -i\int_0^x \mathrm{d}t \exp_{\mathcal{T}}\left(i\int_t^s \mathrm{d}\tau(H(\tau)+V(\tau))\right)V(t)|\psi(t)\rangle \ . \tag{4.124}$$

Hence, using the triangle inequality, we have for all $x \geq 0$ that

$$\||\epsilon(x)\rangle\| \leq \int_0^x \mathrm{d}t \left\|\exp_{\mathcal{T}}\left(i\int_t^x \mathrm{d}\tau(H(\tau)+V(\tau))\right)V(t)|\psi(t)\rangle\right\| \tag{4.125}$$

$$\leq \int_0^x \mathrm{d}t\|V(t)\| \ . \tag{4.126}$$

■

Analytic approximation of observables on gapped ground states

We now state the below theorem, which establishes an analogous result to Theorem 4.3.4 in the case of gapped ground states: for weakly perturbed Hamiltonians, ground state properties depend approximately analytically on the perturbed parameter.

Theorem 4.3.13. *Let $H \in \mathrm{Herm}(\mathcal{H})$ be a Hamiltonian and let $V : \mathbb{C} \to \mathrm{Lin}(\mathcal{H})$ be an analytic family of perturbations such that $H+V(x)$ has a unique ground state $|\psi_0(x)\rangle$ with spectral gap $\gamma > 0$ for all $x \in [-R, R]$, for some $R > 0$. Fix an observable $O \in \mathrm{Herm}(\mathcal{H})$. Then for every $T > 0$, the function*

$$f_{\mathsf{ground}}(x) := \langle\psi_0(x)|O|\psi_0(x)\rangle \tag{4.127}$$

has a (δ, M, R)-analytic approximation where

$$\delta = 2\|O\|c''\lambda R\exp\left(-a''\frac{T}{\log^2(T)}\right) \ , \quad M = \|O\|\exp\left(2RT\lambda e^{2RT\lambda}\right) \ , \tag{4.128}$$

where $a'', c'' > 0$ are constants and $\lambda := \sup_{|z|\leq R}\|\partial_z V(z)\|$.

Proof of Theorem 4.3.13. We can use the spectral flow as prescribed in Lemmas 4.2.9 and 4.2.10 to write

$$\frac{\mathrm{d}}{\mathrm{d}x}|\psi_0(x)\rangle = i\Psi_{H+V(x)}(\partial_x V(x))|\psi_0(x)\rangle \ . \tag{4.129}$$

We define $\tilde{\Psi}_{H+V(x)}(\partial_x V(x))$ to be the spectral flow operator with the integral truncated to the range $[-T, T]$, as in Lemma 4.3.10. Then $\tilde{\Psi}_{H+V(x)}(\partial_x V(x))$ extends to an analytic function for $|z| \leq R$, bounded by

$$\left\| \tilde{\Psi}_{H+V(z)}(\partial_z V(z)) \right\| \leq 2T \|\partial_z V(z)\| \exp\left(2T\|V(z)\|\right) \leq 2T\lambda e^{2TR\lambda} \ , \tag{4.130}$$

where we defined $\lambda := \sup_{|z|\leq R} \|\partial_z V(z)\|$. Furthermore, for $x \in [-R, R]$ we have

$$\left\| \tilde{\Psi}_{H+V(x)}(\partial_x V(x)) - \Psi_{H+V(x)}(\partial_x V(x)) \right\| \leq c'' \lambda \exp\left(-a'' \frac{T}{\log^2(T)} \right) \ , \tag{4.131}$$

where we have absorbed the constants from Lemma 4.3.10 into the constants $c'', a'' > 0$.

Defining $|\tilde{\psi}_0(z)\rangle$ for $|z| \leq R$ by

$$|\tilde{\psi}_0(0)\rangle = |\psi_0(0)\rangle \ , \quad \frac{\mathrm{d}}{\mathrm{d}z} |\tilde{\psi}_0(z)\rangle = i \tilde{\Psi}_{H+V(z)}(\partial_z V(z)) \ , \tag{4.132}$$

we have, by Lemma 4.3.12, that

$$\left\| |\psi_0(x)\rangle - |\tilde{\psi}_0(x)\rangle \right\| \leq c'' \lambda R \exp\left(-a'' \frac{T}{\log^2(T)} \right) \quad \text{for} \quad x \in [-R, R] \ . \tag{4.133}$$

Moreover, by Lemma 4.3.11, we are guaranteed that $|\tilde{\psi}_0(x)\rangle$ extends to an analytic function for $|z| \leq R$ with

$$\sup_{|z|\leq R} \| |\tilde{\psi}_0(z)\rangle \| \leq \exp\left(2RT\lambda e^{2RT\lambda} \right) \ . \tag{4.134}$$

Defining the analytic function $\tilde{f}_{\mathsf{ground}}(z) := \langle \tilde{\psi}_0(z) | O | \tilde{\psi}_0(z) \rangle$, we have

$$\left| f_{\mathsf{ground}}(x) - \tilde{f}_{\mathsf{ground}}(x) \right| \leq 2\|O\| \cdot \left\| |\psi_0(x)\rangle - |\tilde{\psi}_0(x)\rangle \right\| \tag{4.135}$$

$$\leq 2\|O\| c'' \lambda R \exp\left(-a'' \frac{T}{\log^2(T)} \right) \quad \text{for} \quad x \in [-R, R] \ , \tag{4.136}$$

and

$$|f_{\mathsf{ground}}(z)| \leq \|O\| \| |\tilde{\psi}_0(z)\rangle \|^2 \tag{4.137}$$

$$\leq \|O\| \exp\left(4RT\lambda e^{2RT\lambda} \right) \quad \text{for} \quad |z| \leq R \ . \tag{4.138}$$

Hence $\tilde{f}_{\mathsf{ground}}$ is a (δ, M, R)-analytic approximation for f_{ground}, where

$$\delta = 2\|O\| c'' \lambda R \exp\left(-a'' \frac{T}{\log^2(T)} \right) \ , \quad M = \|O\| \exp\left(2RT\lambda e^{2RT\lambda} \right) \ . \tag{4.139}$$

■

We restate the ground state part of Theorem 4.3.2 below:

Theorem 4.3.14 (Restatement of Theorem 4.3.2(II))**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ be a family of Hamiltonians satisfying Assumption 4.3.1*(II)*. Then, for any $\delta > 0$, the function $f_{\mathsf{ground}}(x)$ defined in Eq.* (4.68) *has a (δ, M, R)-analytic approximation, where*

$$M = 16\|O_A\| \ , \quad x_* \geq R = 1/\,\mathcal{O}(\log^{D+1}(\delta^{-1})) \ . \tag{4.140}$$

Proof of Theorem 4.3.14. Let $r_0, T > 0$ be positive parameters which will later be fixed. Similarly to the proof of Theorem 4.3.9, we write $f_{\mathsf{ground}}(x, y; B_{r_0})$ for the function

$$f_{\mathsf{ground}}(x, y; B_{r_0}(A)^c) := \langle \psi_0(x, y; B_{r_0}(A)^c)|O_A|\psi_0(x, y; B_{r_0}(A)^c)\rangle \ , \tag{4.141}$$

where $|\psi_0(x, y; B_{r_0}(A)^c)\rangle$ is defined, as in Assumption 4.3.1(II), as the unique ground state state corresponding to the Hamiltonian

$$H(x, y; B_{r_0}(A)^c) := \sum_{A'\subseteq B_{r_0}(A)^c} H_{A'}(x) + \sum_{\substack{A'\subseteq\Gamma \\ A'\nsubseteq B_{r_0}(A)^c}} H_{A'}(y) \ . \tag{4.142}$$

By the GALI property for Ground states (Lemma 4.2.17), $f_{\mathsf{ground}}(x, y; B_{r_0}(A)^c)$ only depends very slightly on its first argument, in the sense that we can bound

$$|f_{\mathsf{ground}}(x, x; B_{r_0}(A)^c) - f_{\mathsf{ground}}(0, x; B_{r_0}(A)^c)| \leq |x|c_2\|O_A\||A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_r a(r-1)/\log^2(a(r-1))} \tag{4.143}$$

$$\leq c'|x|e^{-a' r_0/\log^2(r_0)} \ , \tag{4.144}$$

for some constants $b_2, c_2 > 0$, and where on the second line we have absorbed all the constants in the expression into some $a', c' > 0$. As in the proof of Theorem 4.3.9, we can bound

$$\|\partial_x H(0, x; B_{r_0}(A)^c)\| \leq \sum_{\substack{A'\subseteq\Gamma \\ A'\nsubseteq B_{r_0}(A)^c}} \|\partial_x H_{A'}(x)\| \tag{4.145}$$

$$\leq \sum_{i\in B_{r_0}(A)} \sum_{\substack{A'\subseteq\Gamma \\ i\in A'}} \|\partial_x H_{A'}(x)\| \tag{4.146}$$

$$\leq |B_{r_0}(A)|\|\partial_x H(x)\|_{F_g}\|F_g\| \tag{4.147}$$

$$\leq b' r_0^D \ , \tag{4.148}$$

for some constant $b' > 0$. Thus, by Theorem 4.3.13, there exists a (δ', M, x_*) approximation $\tilde{f}_{\mathsf{ground}}(x)$ for $f_{\mathsf{ground}}(0, x; B_{r_0}(A)^c)$, where

$$\delta' = 2\|O_A\|c''b' r_0^D R \exp\left(-a''\frac{T}{\log^2 T}\right) \ , \quad M = \|O_A\| \exp\left(2x_* T b' r_0^D e^{2RTb' r_0^D}\right) \ . \tag{4.149}$$

Hence, by Eq. (4.144) $\tilde{f}_{\mathsf{ground}}(x)$ is a (δ, M, x_*)-approximation for $f_{\mathsf{ground}}(x)$, where

$$\delta = c' x_* e^{-a' r_0/\log^2(r_0)} + 2\|O_A\| c'' b' r_0^D x_* \exp\left(-a'' \frac{T}{\log^2 T}\right) \ , \quad M = \|O_A\| \exp\left(2 x_* T b' r_0^D e^{2RTb' r_0^D}\right) \ . \tag{4.150}$$

Note that we have the freedom to tune the parameters $r_0, T > 0$, as well as the ability to restrict to a smaller $R \leq x_*$. We thus set $T = r_0$, and choose $R = \min\{x_*, 1/(2Tb' r_0^D)\}$. This leads to

$$\delta = r_0^{-1} \exp\left(-\Omega(r_0/\log^2(r_0))\right) \ , \quad M \leq \|O_A\| e^e \leq 16\|O_A\| \ . \tag{4.151}$$

Hence any arbitrarily small $\delta > 0$ can be achieved by setting $r_0 \sim \log(\delta^{-1})$, giving

$$R = 1/\,\mathcal{O}(r_0^{D+1}) = 1/\,\mathcal{O}(\log^{D+1}(\delta^{-1})) \ , \tag{4.152}$$

as required. ∎

4.4 Perturbation theory and simulation gadgets

4.4.1 Effective Hamiltonians from local gaps

In the analogue quantum simulation literature (for example, Refs. [72, 135, 86, 56]), perturbative "gadgets" are used to simulate complicated local Hamiltonian terms using simpler building blocks. These work by introducing ancillary qubits to mediate interactions, with large on-site energy penalty terms forcing the ancillary qubits into the correct subspace described by some projector $\mathbb{1} \otimes P_0$ (where P_0 acts only on the ancillary qubits). Provided that the local energy penalties are high enough (i.e., scaling linearly with the size of the system), this defines a low-energy subspace of the global Hamiltonian in which an effective Hamiltonian acting on the non-ancillary qubits can be derived.

In this section we establish the perturbation theory tools required to deal with gadget Hamiltonians in the weak-interaction regime, where the strength of the local penalty terms is constant and hence a global energy gap is not present, but where we can still meaningfully define a perturbative effective Hamiltonian induced by the gadgets. We will make use of the local Schrieffer-Wolff transformation, for which the basic idea is as follows. Given a Hamiltonian H' consisting of many gadgets with relatively weak interactions, we cannot guarantee that the global low-energy subspace will coincide with the "all gadget ancillas in ground state" subspace, which is described by some projector $\mathbb{1} \otimes P_0$ acting on the ancillary qubits, even though the gadgets may be locally gapped. However, the subspace described by $\mathbb{1} \otimes P_0$ can be slightly rotated to give a subspace which *is* invariant under H' — though this does not necessarily correspond to a subspace spanned only by its lowest-energy eigenvectors. The local Schrieffer-Wolff transformation provides a generator T, which T/i is a local Hamiltonian, which generates this rotation from the invariant subspace into the $\mathbb{1} \otimes P_0$ space, so that we can define an effective Hamiltonian $H_{\mathsf{eff}} \otimes P_0 = (\mathbb{1} \otimes P_0) e^T H' e^{-T} (\mathbb{1} \otimes P_0)$. This effective Hamiltonian

corresponds to the space in which the gadgets are behaving as designed, and in which the usual effective Hamiltonian is recovered in the limit of large interaction strengths.

The construction of the local Schrieffer-Wolff transformation, originally introduced by Ref. [206], is explained in Section 4.4.2, and various basic properties are proven (with several proofs deferred to Appendix 4.6.3). Our analysis closely follows the excellent review Ref. [172], but differs in two respects. Firstly, we will derive bounds on the effective Hamiltonian in terms of the $\|\cdot\|_F$-norm for a normalised F function F — this will be useful when we later analyse dynamics under such Hamiltonians, and in fact turns out to lead to slightly simpler analysis than the norm used in Ref. [172], in particular leading to a convergent power series which does not need to be truncated at finite order. Secondly, we generalise to the case where the simulator Hamiltonian is a polynomial in the perturbative parameter, allowing interaction strengths over several energy scales — this is the relevant situation for perturbative gadgets. We define such gadgets formally in Section 4.4.3 and prove several useful properties in Appendix 4.6.4.

4.4.2 Local Schrieffer-Wolff perturbation theory

Consider a Hilbert space $\mathcal{H} = \bigotimes_{i\in\Gamma}\mathcal{H}_i$ consisting of $|\Gamma| = n$ sites. We will study Hamiltonians of the form

$$H(x) = \sum_{\alpha=0}^{d} x^\alpha H^{(\alpha)} \ , \tag{4.153}$$

parametrised by some real number $x \geq 0$.[1] The constant term $H^{(0)}$ is assumed to be 1-local with local ground state projectors (which may be the identity) $P_{0,i}$ at every $i \in \Gamma$, such that when a locally excited space exists, it is separated by a gap of at least $\Delta > 0$ above the ground space which is assumed to have zero energy. The higher-order terms are assumed to have interactions bounded by

$$\|H^{(\alpha)}\|_F \leq J \ , \quad 1 \leq \alpha \leq d \ , \tag{4.154}$$

for some normalised F-function F, and constant $J > 0$. Our goal is to construct a power series

$$T(x) = \sum_{q=1}^{\infty} x^q T^{(q)} \ , \tag{4.155}$$

and some series $\{V^{(q)}\}_{q\geq 1}$ of Hamiltonians such that

$$e^{T(x)} H(x) e^{-T(x)} = H^{(0)} + \sum_{q=1}^{\infty} x^q \left(V^{(q)} + [T^{(q)}, H^{(0)}] \right) \ , \tag{4.156}$$

[1] Ultimately we will be interested in the object $x^{-d}H(x)$, containing strong interactions which become singular as $x \to 0$.

and where, for $q \geq 1$, the term $V^{(q)} + [T^{(q)}, H^{(0)}]$ is local and block-diagonal with respect to the projector $P_0 := \bigotimes_{i\in\Gamma} P_{0,i}$. In other words, we wish to locally diagonalise $H(x)$ with respect to this P_0 (for x sufficiently small that the series converges). Following Ref. [172], the strategy will be to inductively construct $\{V^{(q)}\}_{q\geq 1}$ and $\{T^{(q)}\}_{q\geq 1}$ as follows:

- The first term $V^{(1)}$ is set to $V^{(1)} = H^{(1)}$.
- For $q \geq 1$, $T^{(q)}$ is defined in terms of $V^{(q)}$.
- For $q \geq 2$, $V^{(q)}$ is defined in terms of the lower-order $T^{(q-1)}, T^{(q-2)}, \ldots$.

Note that Eq. (4.156) fully determines $V^{(q)}$ in terms of the lower-order $T^{(q)}$ by comparing powers of x^q: we can expand the left-hand side in powers of x to find (defining $H^{(\alpha)} = 0$ for $\alpha > d$):

$$e^{T(x)}H(x)e^{-T(x)} = \sum_{\alpha=0}^{\infty} x^\alpha e^{T(x)} H^{(\alpha)} e^{-T(x)} \tag{4.157}$$

$$= \sum_{\alpha=0}^{\infty} x^\alpha \left(H^{(\alpha)} + \sum_{r=1}^{\infty} \frac{1}{r!} \sum_{q_1,\ldots,q_r=1}^{\infty} x^{q_1+\cdots+q_r} [T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(\alpha)}] \cdots]] \right) \tag{4.158}$$

$$\begin{aligned} &= H^{(0)} + \sum_{q=1}^{\infty} x^q \Bigg([T^{(q)}, H^{(0)}] + H^{(q)} + \\ &\quad + \sum_{r=2}^{q} \frac{1}{r!} \sum_{\substack{q-1\geq q_1,\ldots,q_r\geq 1 \\ q_1+\cdots+q_r=q}} [T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(0)}] \cdots]] \\ &\quad + \sum_{\alpha=1}^{q} \sum_{r=1}^{q-\alpha} \frac{1}{r!} \sum_{\substack{q-\alpha\geq q_1,\ldots,q_r\geq 1 \\ q_1+\cdots+q_r=q-\alpha}} [T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(\alpha)}] \cdots]] \Bigg) , \end{aligned} \tag{4.159}$$

where in the third line we have separated the sum over r into the cases $\alpha = 0$ and $\alpha \geq 1$, and further separated the former term into the $r = 1$ case and the $r \geq 2$ case. Hence

$$V^{(1)} = H^{(1)} , \tag{4.160}$$

$$\begin{aligned} V^{(q)} = H^{(q)} &+ \sum_{r=2}^{q} \frac{1}{r!} \sum_{\substack{q-1\geq q_1,\ldots,q_r\geq 1 \\ q_1+\cdots+q_r=q}} [T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(0)}] \cdots]] \\ &+ \sum_{\alpha=1}^{q} \sum_{r=1}^{q-\alpha} \frac{1}{r!} \sum_{\substack{q-\alpha\geq q_1,\ldots,q_r\geq 1 \\ q_1+\cdots+q_r=q-\alpha}} [T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(\alpha)}] \cdots]] , \quad q \geq 2 . \end{aligned} \tag{4.161}$$

Note that the expression for $V^{(q)}$ only depends on $T^{(q')}$ for $q' < q$ (and ultimately $T^{(q')}$ will depend only on $V^{(q')}$), so this inductive definition is well-defined. Note here that we define $H^{(q)} = 0$ for $q > d$. We only have the freedom to choose $T^{(q)}$ to satisfy the desired block-diagonal condition, which we do in the following.

Given $A \subseteq \Gamma$, we define the projectors

$$P_A := \bigotimes_{i \in A} P_{0,i} \ , \quad Q_A := 1 - P_A \ . \tag{4.162}$$

Note that here, and below, we abuse notation and suppress tensor product factors of the identity on the non-ancillary qubits (that is, identifying P_A with $P_A \otimes \mathbb{1}_{\Gamma \setminus A}$). For any operator X, we define its off-diagonal and diagonal parts with respect to this pair of projectors by

$$\mathcal{O}_A(X) := P_A X Q_A + Q_A X P_A \ , \tag{4.163}$$

$$\mathcal{D}_A(X) := P_A X P_A + Q_A X Q_A \ . \tag{4.164}$$

We now define the superoperator $\mathcal{L}_A$ by

$$\mathcal{L}_A(X) := \left(\int_0^\infty \mathrm{d}t e^{-tH^{(0)}|_A/\Delta} \right) Q_A X P_A - P_A X Q_A \left(\int_0^\infty \mathrm{d}t e^{-tH^{(0)}|_A/\Delta} \right) \ , \tag{4.165}$$

where $H^{(0)}|_A$ consists of the 1-local terms in $H^{(0)}$ acting on the sites $i \in A$, i.e.

$$H^{(0)}|_A := \sum_{i \in A} H_i^{(0)} \ , \tag{4.166}$$

and $\Delta > 0$ is the size of the local energy gaps in $H^{(0)}$. This satisfies the following properties:

Lemma 4.4.1 (Properties of $\mathcal{L}_A$ — see [172] Section 4.4)**.** *Let the superoperator $\mathcal{L}_A$ be defined as in Eq.* (4.165)*.*

(I) For any X, we have

$$\| \mathcal{L}_A(X) \| \leq \|X\| \ . \tag{4.167}$$

(II) If X acts only on the spins $A \subseteq \Gamma$, then so does $\mathcal{L}_A(X)$.

(III) If X acts only on the spins A, then

$$\mathcal{L}_A\left([H^{(0)}, X] \right) = \left[H^{(0)}, \mathcal{L}_A(X) \right] = \Delta \, \mathcal{O}_A(X) \ . \tag{4.168}$$

Proof of Lemma 4.4.1. Point (I) follows from the definition of Q_A, which projects onto the excited space of $H^{(0)}|_A$ (with energy at least Δ) and thus guarantees that $\|e^{-tH^{(0)}|_A/\Delta}Q_A\| \leq e^{-t}$.

Point (II) follows immediately from the definition, as X, Q_A, and P_A are all supported only on A.

For point (III), note that we have

$$Q_A[H^{(0)}, X]P_A = Q_A[H^{(0)}|_A, X]P_A \tag{4.169}$$

$$= H^{(0)}|_A Q_A X P_A \ , \tag{4.170}$$

where the first equality follows as X is only supported on A, and the second equality follows as $H^{(0)}|_A P_A = 0$, and $H^{(0)}|_A$ commutes with Q_A. Thus

$$\left(\int_0^\infty \mathrm{d}t e^{-tH^{(0)}|_A/\Delta}\right) Q_A[H^{(0)}, X]P_A = \left(\int_0^\infty \mathrm{d}t e^{-tH^{(0)}|_A/\Delta}\right) H^{(0)}|_A Q_A X P_A \tag{4.171}$$

$$= \Delta Q_A X P_A \ . \tag{4.172}$$

Hence, via an identical calculation for the second term,

$$\mathcal{L}_A\left([H^{(0)}, X]\right) = Q_A X_A P_A + P_A X_A Q_A \ , \tag{4.173}$$

as required. The fact that this is equivalent to $[H^{(0)}|_A, \mathcal{L}_A(X)] = [H^{(0)}, \mathcal{L}_A(X)]$ follows from a similar computation, by commuting $H^{(0)}|_A$ to the outside of both terms. ∎

Writing $V^{(q)}$ as a sum of local terms

$$V^{(q)} = \sum_{A \subseteq \Gamma} V_A^{(q)} \ , \tag{4.174}$$

we define $T^{(q)}$ by

$$T^{(q)} = \sum_{A \subseteq \Gamma} \Delta^{-1}\, \mathcal{L}_A(V_A^{(q)}) \ . \tag{4.175}$$

Observe that by Lemma 4.4.1, $T^{(q)}$ inherits the locality properties from $V^{(q)}$, and

$$[T^{(q)}, H^{(0)}] + V^{(q)} = \sum_{A \subseteq \Gamma} \left(\Delta^{-1}[\mathcal{L}_A(V_A^{(q)}), H^{(0)}] + V_A^{(q)}\right) \tag{4.176}$$

$$= \sum_{A \subseteq \Gamma} \left(-\mathcal{O}_A(V_A^{(q)}) + V_A^{(q)}\right) \tag{4.177}$$

$$= \sum_{A \subseteq \Gamma} \mathcal{D}_A(V_A^{(q)}) \ , \tag{4.178}$$

where for the second equality we used Lemma 4.4.1(III). Hence the right-hand side of Eq. (4.156) is indeed a sum of local and block-diagonal interactions. We will generally denote $V(x) = \sum_{q \geq 1} x^q V^{(q)}$, and write $V_{\mathsf{eff}}(x)$ for the effective Hamiltonian in the $\mathbb{1} \otimes P_0$ space, that is

$$V_{\mathsf{eff}}(x) \otimes P_0 := \sum_{q \geq 1} x^q (\mathbb{1} \otimes P_0) V^{(q)} (\mathbb{1} \otimes P_0) \ . \tag{4.179}$$

Bounding the perturbative expansion

Having obtained a recursive definition for the perturbative series $V^{(q)}$ and $T^{(q)}$ in Eq. (4.156), we can prove bounds on the growth of these operators with q, ultimately establishing that the effective Hamiltonian is itself bounded in the $\|\cdot\|_F$-norm. This main result is stated below, as Lemma 4.4.2, and is proved in Appendix 4.6.3.

Lemma 4.4.2 (Compare to Ref. [172], Lemma 4.2 — see Lemma 4.6.9)**.** *For $q \geq 1$, $V^{(q)}$ is bounded as*

$$\|V^{(q)}\|_F \leq \frac{\Delta\theta^q}{16} \ , \tag{4.180}$$

where $\theta > 0$ is a constant depending on the ratio J/Δ. In particular, if $x \leq 1/(2\theta)$, then the Hamiltonian $V(x) = \sum_{q\geq 1} x^q V^{(q)}$ is bounded as

$$\|V_{\mathsf{eff}}(x)\|_F \leq \|V(x)\|_F \leq \frac{\Delta\theta x}{8} \ . \tag{4.181}$$

From Lemma 4.4.2, we can establish the following bound on the generator $T(x)$.

Corollary 4.4.3. *For $q \geq 1$, $T^{(q)}$ is bounded as*

$$\|T^{(q)}\|_F \leq \frac{\theta^q}{16} \ , \tag{4.182}$$

where $\theta > 0$ is the constant from Lemma 4.4.2. Moreover, if $x \leq 1/(2\theta)$, then the generator of the Schrieffer-Wolff transformation $T(x) = \sum_{q\geq 1} x^q T^{(q)}$ is bounded as

$$\|T(x)\|_F \leq \frac{\theta x}{8} \ . \tag{4.183}$$

Proof of Corollary 4.4.3. Follows immediately from Eq. (4.175) and Lemma 4.4.1(I), using the bound on $\|V^{(q)}\|_F$ from Lemma 4.4.2. ∎

Gibbs states and ground states of the perturbed Hamiltonian

As mentioned, the subspace $\mathbb{1}\otimes P_0$ will not necessarily contain only the lowest-energy states of the rotated Hamiltonian $e^{T(x)}H(x)e^{-T(x)}$, unless x is very small — typically $x \sim n^{-1}$, where $n = |\Gamma|$, is necessary to produce the necessary gap in the spectrum. In this section we establish that, even for constant x, low-temperature Gibbs states have high overlap within the $\mathbb{1}\otimes P_0$ subspace and are thus well-approximated by the Gibbs states of the effective Hamiltonian. The approximation improves as the inverse-temperature increases, and becomes exact for ground states (see Corollary 4.4.6).

The following result, Lemma 4.4.4, is the key tool for this purpose. This lemma shows that a Hamiltonian of the form $H^{(0)} + V$, where $H^{(0)}$ consists of gapped 1-local terms and V is geometrically local (and which is locally block-diagonal with respect to

the terms in $H^{(0)}$), has Gibbs states which are likely to be in the ground space of $H^{(0)}$. Intuitively, this is because any excitation of l terms in $H^{(0)}$ will be suppressed as $e^{-\beta l}$ in the Gibbs state of inverse temperature β compared to its non-excited counterpart. Via some combinatorics to account for the terms at different levels of excitation, we can thus show that the non-excited space is strongly favoured in the Gibbs state.

Lemma 4.4.4. *Let $H^{(0)}$ and V be Hamiltonians on $\mathcal{H} = \bigotimes_{i\in\Gamma} \mathcal{H}_i$, where $\mathcal{H}_i \cong \mathbb{C}^m$, with the following properties:*

(a) There is a subset $\Gamma_{\mathsf{anc}} \subseteq \Gamma$ of "ancillary sites" such that $H^{(0)}$ takes the form

$$H^{(0)} = \Delta \sum_{i\in\Gamma_{\mathsf{anc}}} (\mathbb{1} - P_{0,i}) \ , \tag{4.184}$$

where, for every $i \in \Gamma_{\mathsf{anc}}$, $P_{0,i}$ is a 1-local projector acting on site i, of rank 1. We write $P_0 = \bigotimes_{i\in\Gamma_{\mathsf{anc}}} P_{0,i}$ for the projector onto the ground space of $H^{(0)}$.

(b) $V = \sum_{A\subseteq\Gamma} V_A$ is local on Γ, with bounded $\|\cdot\|_F$ norm for some F-function F. Moreover, V_A is block-diagonal with respect to $P_{0,A} := \bigotimes_{i\in A\cap\Gamma_{\mathsf{anc}}} P_{0,i}$ for every $A \subseteq \Gamma$.

For any $\beta \geq 0$, we write σ_β for the Gibbs state of the Hamiltonian $H^{(0)} + V$, i.e.,

$$\sigma_\beta := \frac{e^{-\beta(H^{(0)}+V)}}{\mathrm{tr}[e^{-\beta(H^{(0)}+V)}]} \ . \tag{4.185}$$

Then

$$\mathrm{tr}[P_0\sigma_\beta] \geq \exp\left(-(m-1)|\Gamma_{\mathsf{anc}}|e^{-\beta(\Delta-3\|V\|_F\|F\|)}\right) \ . \tag{4.186}$$

This result only gives a non-trivial bound in the very low-temperature regime $\beta \gtrsim \log|\Gamma_{\mathsf{anc}}| \sim \log(n)$, however in our analysis of gadget Hamiltonians we will ultimately absorb a factor of x^{-d} into β before applying this theorem to show that the Gibbs state approximately lies in the effective subspace P_0 for $x \sim 1/\log n$.

Proof of Lemma 4.4.4. Since $[V, P_0] = 0$, it follows that V is block-diagonal with respect to the projectors $P_0, 1-P_0$. We write $\{|\psi_a\rangle\}_a$ for the eigenvectors of P_0VP_0, i.e. an orthornormal set spanning the image of P_0, such that $P_0|\psi_a\rangle = |\psi_a\rangle$, $V|\psi_a\rangle = E_a|\psi_a\rangle$ for some energies $E_a \in \mathbb{R}$.

Since each $P_{0,i}$ is a rank 1 projector, we can without loss of generality choose a basis $\{|b\rangle\}_{b=0}^{m-1}$ of each site $\mathcal{H}_i$, such that $P_{0,i} = |0\rangle\langle 0|_i$. We define the 1-local unitary X to be such that $X|b\rangle = |b+1 \mod m\rangle$. For $\mathbf{b} \in \{0,\ldots,m\}^{\Gamma_{\mathsf{anc}}}$, we write $X^{\mathbf{b}}$ for the tensor product

$$X^{\mathbf{b}} := \bigotimes_{i\in\Gamma_{\mathsf{anc}}} X_i^{b_i} \ , \tag{4.187}$$

which shifts the basis element at each ancilla site i by b_i. Since the $\{|\psi_a\rangle\}_a$ span the entire space in which all the ancillas are in the $|0\rangle$ state, it follows that $\{|\psi_{a,\mathbf{b}}\rangle := X^{\mathbf{b}}|\psi_a\rangle\}_{a,\mathbf{b}}$ is a basis for the entire space $\mathcal{H}$. Hence we can write

$$\mathrm{tr}\left[e^{-\beta(H^{(0)}+V)}\right] = \sum_{a,\mathbf{b}} \langle\psi_a|(X^{\mathbf{b}})^\dagger e^{-\beta(H^{(0)}+V)} X^{\mathbf{b}}|\psi_a\rangle \tag{4.188}$$

$$= \sum_{a,\mathbf{b}} \langle\psi_a|e^{-\beta(X^{\mathbf{b}})^\dagger(H^{(0)}+V)X^{\mathbf{b}}}|\psi_a\rangle \ . \tag{4.189}$$

We now decompose the exponentiated Hamiltonian into its blocks with respect to the projectors P_0 and $(1-P_0)$:

$$(X^{\mathbf{b}})^\dagger(H^{(0)}+V)X^{\mathbf{b}} = P_0(X^{\mathbf{b}})^\dagger(H^{(0)}+V)X^{\mathbf{b}}P_0 \tag{4.190}$$

$$+ (1-P_0)(X^{\mathbf{b}})^\dagger(H^{(0)}+V)X^{\mathbf{b}}(1-P_0) \tag{4.191}$$

$$P_0(X^{\mathbf{b}})^\dagger(H^{(0)}+V)X^{\mathbf{b}}(1-P_0) + \text{h.c.} \ . \tag{4.192}$$

Notice that, since $X^{\mathbf{b}}$ maps a zero-energy state $|0\rangle_i$ to an excited state $|b_i\rangle_i$ for every $b_i \neq 0$, we have

$$P_0(X^{\mathbf{b}})^\dagger H^{(0)} X^{\mathbf{b}} P_0 = \Delta|\mathbf{b}|P_0 \ , \tag{4.193}$$

where $|\mathbf{b}|$ denotes the Hamming weight of $\mathbf{b}$. Moreover, we can bound

$$\left\|P_0(X^{\mathbf{b}})^\dagger V X^{\mathbf{b}} P_0 - P_0 V P_0\right\| \leq \sum_{A\subseteq\Gamma} \left\|P_0(X^{\mathbf{b}})^\dagger V_A X^{\mathbf{b}} P_0 - P_0 V_A P_0\right\| \tag{4.194}$$

$$\leq \sum_{\substack{i\in\Gamma_{\mathsf{anc}}\\ b_i\neq 0}} \sum_{\substack{A\subseteq\Gamma\\ i\in A}} \left\|P_0(X^{\mathbf{b}})^\dagger V_A X^{\mathbf{b}} P_0 - P_0 V_A P_0\right\| \tag{4.195}$$

$$\leq \sum_{\substack{i\in\Gamma_{\mathsf{anc}}\\ b_i\neq 0}} \sum_{\substack{A\subseteq\Gamma\\ i\in A}} 2\|V_A\| \tag{4.196}$$

$$\leq 2|\mathbf{b}|\|F\|\|V\|_F \ , \tag{4.197}$$

where in the first line we used the triangle inequality, in the second line we used the fact that $X^{\mathbf{b}}$ and V_A commute if $b_i = 0$ for all $i \in A$. Since $(X^{\mathbf{b}})^\dagger H^{(0)} X^{\mathbf{b}}$ is block-diagonal with respect to P_0, we have

$$P_0(X^{\mathbf{b}})^\dagger H^{(0)} X^{\mathbf{b}}(1-P_0) = 0 \ . \tag{4.198}$$

Finally, using that V is locally block-diagonal with respect to the $P_{0,A}$, and again that $X^{\mathbf{b}}$ is only supported on i with $b_i \neq 0$, we have

$$P_0(X^{\mathbf{b}})^\dagger V X^{\mathbf{b}}(1-P_0) \leq \sum_{A\subseteq\Gamma} \left\|P_0(X^{\mathbf{b}})^\dagger V_A X^{\mathbf{b}}(1-P_0)\right\| \tag{4.199}$$

$$\leq \sum_{\substack{i\in\Gamma\\ b_i\neq 0}} \sum_{\substack{A\subseteq\Gamma\\ i\in A}} \|V_A\| \tag{4.200}$$

$$\leq |\mathbf{b}|\|F\|\|V\|_F \ . \tag{4.201}$$

Putting these together, we can deduce that

$$(X^{\mathbf{b}})^{\dagger}(H^{(0)}+V)X^{\mathbf{b}} \geq \Delta|\mathbf{b}|P_0 + P_0VP_0 - 3|\mathbf{b}|\|F\|\|V\|_F P_0 \tag{4.202}$$

$$+ (1-P_0)(X^{\mathbf{b}})^{\dagger}(H^{(0)}+V)X^{\mathbf{b}}(1-P_0) \ , \tag{4.203}$$

and hence

$$\mathrm{tr}\left[e^{-\beta(H^{(0)}+V)}\right] \leq \sum_{a,\mathbf{b}} \langle\psi_a|e^{-\beta(\Delta|\mathbf{b}|P_0+P_0VP_0-3|\mathbf{b}|\|F\|\|V\|_F P_0)}|\psi_a\rangle \tag{4.204}$$

$$= \sum_{a,\mathbf{b}} e^{-\beta(\Delta|\mathbf{b}|+E_a-3|\mathbf{b}|\|F\|\|V\|_F)} \ , \tag{4.205}$$

where we ignored the $(1-P_0)$-block, since the $|\psi_a\rangle$ are in the span of P_0. Notice that

$$\mathrm{tr}\left[P_0 e^{-\beta(H^{(0)}+V)}\right] = \sum_a e^{-\beta E_a} \ , \tag{4.206}$$

and hence we can write

$$\mathrm{tr}\left[e^{-\beta(H^{(0)}+V)}\right] \leq \mathrm{tr}\left[P_0 e^{-\beta(H^{(0)}+V)}\right] \sum_{\mathbf{b}\in\{0,\ldots,m-1\}^{\Gamma_{\mathsf{anc}}}} e^{-\beta|\mathbf{b}|(\Delta-3\|F\|\|V\|_F)} \ , \tag{4.207}$$

which can be rearranged to give

$$\frac{1}{\mathrm{tr}\left[P_0\sigma_\beta\right]} \leq \sum_{l=0}^{|\Gamma_{\mathsf{anc}}|} \sum_{\substack{\mathbf{b}\in\{0,\ldots,m-1\}^{\Gamma_{\mathsf{anc}}} \\ |\mathbf{b}|=l}} e^{-\beta l(\Delta-3\|F\|\|V\|_F)} \tag{4.208}$$

$$= \sum_{l=0}^{|\Gamma_{\mathsf{anc}}|} \binom{|\Gamma_{\mathsf{anc}}|}{l} (m-1)^l e^{-\beta l(\Delta-3\|F\|\|V\|_F)} \tag{4.209}$$

$$= \left(1+(m-1)e^{-\beta(\Delta-3\|F\|\|V\|_F)}\right)^{|\Gamma_{\mathsf{anc}}|} \tag{4.210}$$

$$\leq \exp\left((m-1)|\Gamma_{\mathsf{anc}}|e^{-\beta(\Delta-3\|F\|\|V\|_F)}\right) \ , \tag{4.211}$$

from which the result follows. ∎

We now specialise this result to the setting of local Schrieffer-Wolff perturbation theory, showing that low-temperature Gibbs states of the perturbed Hamiltonian approximately lie within the subspace of the effective Hamiltonian.

Corollary 4.4.5. *Let $H(x)$ be a Hamiltonian as defined in Eqs.* (4.153)-(4.156)*, where the constant term $H^{(0)}$ takes the form given in Eq.* (4.184)*. Let $T(x)$ and $V(x)$ denote the local Schrieffer-Wolff transformation and associated perturbative series. Let $\rho_\beta(x)$*

be the Gibbs state associated to $H(x)$, and let $\rho_{\beta,\mathsf{eff}}(x)$ be the Gibbs state associated to $V_{\mathsf{eff}}(x)$ on the subspace $\mathbb{1}\otimes P_0$ as defined in Eq. (4.179)*, that is*

$$\rho_\beta(x) := \frac{e^{-\beta H(x)}}{\mathrm{tr}[e^{-\beta H(x)}]} \ , \quad \rho_{\beta,\mathsf{eff}}(x) := \frac{e^{-\beta V_{\mathsf{eff}}(x)}}{\mathrm{tr}[e^{-\beta V_{\mathsf{eff}}(x)}]} \ . \tag{4.212}$$

Then, provided $x \leq 1/(2\theta)$ (where θ is the constant from Lemma 4.4.2),

$$\left\| e^{T(x)}\rho_\beta(x)e^{-T(x)} - \rho_{\beta,\mathsf{eff}}(x)\otimes P_0 \right\|_1 \leq 2 - 2\exp\left(-(m-1)|\Gamma_{\mathsf{anc}}|e^{-\beta\Delta(1-3\theta\|F\|x/8)}\right) \ . \tag{4.213}$$

Proof of Corollary 4.4.5. By the definition of the local Schrieffer-Wolff transformation, we have

$$e^{T(x)}H(x)e^{-T(x)} = H^{(0)} + V(x) \ , \tag{4.214}$$

where $H^{(0)}$ and $V(x)$ satisfy the conditions of Lemma 4.4.4, and by Lemma 4.4.2 we have

$$\|V(x)\|_F \leq \frac{\Delta\theta x}{8} \ . \tag{4.215}$$

Hence, by Lemma 4.4.4, we have

$$\mathrm{tr}\left[(\mathbb{1}\otimes P_0)e^{T(x)}\rho_\beta(x)e^{-T(x)}\right] \geq \exp\left(-(m-1)|\Gamma_{\mathsf{anc}}|e^{-\beta\Delta(1-3\theta\|F\|x/8)}\right) =: p_0 \ . \tag{4.216}$$

By definition,

$$\left\| e^{T(x)}\rho_\beta(x)e^{-T(x)} - \rho_{\beta,\mathsf{eff}}(x)\otimes P_0 \right\|_1 = \left\| e^{T(x)}\rho_\beta(x)e^{-T(x)} - p_0^{-1}(\mathbb{1}\otimes P_0)e^{T(x)}\rho_\beta(x)e^{-T(x)}(\mathbb{1}\otimes P_0) \right\|_1 \tag{4.217}$$

$$= \Big\| (\mathbb{1}-\mathbb{1}\otimes P_0)e^{T(x)}\rho_\beta(x)e^{-T(x)}(\mathbb{1}-\mathbb{1}\otimes P_0) + (1-p_0^{-1})(\mathbb{1}\otimes P_0)e^{T(x)}\rho_\beta(x)e^{-T(x)}(\mathbb{1}\otimes P_0) \Big\|_1 \tag{4.218}$$

$$\leq (1-p_0) + |1-p_0^{-1}|p_0 = 2(1-p_0) \ , \tag{4.219}$$

where the final line follows from the triangle inequality. ∎

Taking the limit $\beta \to \infty$ in Corollary 4.4.5, we can obtain an analogous result for ground states. In particular this implies that, when the parameter x is below a constant, all ground states of the perturbed Hamiltonian $H(x)$ lie in the subspace $\mathbb{1}\otimes P_0$.

Corollary 4.4.6 (Compare with Ref. [172], Lemma 4.1)**.** *Let $H(x)$ be a Hamiltonian as defined in Eqs.* (4.153)-(4.156)*, where the constant term $H^{(0)}$ takes the form given in Eq.* (4.184)*. Assume that*

$$x < x_* := \frac{1}{\theta(2 + \|F\|)} \ , \tag{4.220}$$

where θ is the constant from Lemma 4.4.2. Then the ground space of $e^{T(x)}H(x)e^{-T(x)}$ lies in the image of the projector $\mathbb{1} \otimes P_0$.

Proof of Corollary 4.4.6. This follows by taking $\beta \to \infty$ in Corollary 4.4.5. In particular, note that

$$\lim_{\beta\to\infty} e^{T(x)}\rho_\beta(x)e^{-T(x)} = \frac{P(x)}{\mathrm{tr}[P(x)]} \tag{4.221}$$

$$\lim_{\beta\to\infty} \rho_{\beta,\mathsf{eff}}(x) \otimes P_0 = \frac{(\mathbb{1} \otimes P_0)P(x)(\mathbb{1} \otimes P_0)}{\mathrm{tr}[(\mathbb{1} \otimes P_0)P(x)]} \ , \tag{4.222}$$

where $P(x)$ is the ground state projector of $e^{T(x)}H(x)e^{-T(x)}$, whilst Corollary 4.4.5 implies that

$$\lim_{\beta\to\infty} \left\| e^{T(x)}\rho_\beta(x)e^{-T(x)} - \rho_{\beta,\mathsf{eff}}(x) \otimes P_0 \right\| = 0 \ . \tag{4.223}$$

Hence we can conclude that $(\mathbb{1} \otimes P_0)P(x)(\mathbb{1} \otimes P_0) = P(x)$, as required. ∎

4.4.3 Hamiltonian gadgets

In this section, we specialise the local Schrieffer-Wolff transformation constructed in Section 4.4.2 to the case of perturbative gadgets. Generally, these are designed such that the target Hamiltonian H only appears at dth order in the perturbative expansion for $H_{\mathsf{eff}}(x)$ — and hence, the entire Hamiltonian must be rescaled by x^{-d} to make this term constant. This leads to a singular Hamiltonian with interactions polynomial in x^{-1}, but such that $H_{\mathsf{eff}}(x)$ is polynomial in x. We state this explicitly below.

Definition 4.4.7 (Gadget of degree d)**.** Let $\Gamma = \Gamma_{\mathsf{anc}} \cup \Gamma_{\mathsf{eff}}$ be a disjoint partition of the sites Γ, $|\Gamma| = n$. For $d \geq 1$, let $H'(x)$ be a family of Hamiltonians of the form

$$H'(x) = x^{-d} \sum_{\alpha=0}^{d} x^\alpha H^{(\alpha)} \ , \tag{4.224}$$

where $H^{(0)} = \Delta \sum_{i\in\Gamma_{\mathsf{anc}}} (\mathbb{1} - |0_i\rangle\langle 0_i|)$, where $|0_i\rangle$ is a state on $\mathcal{H}_i$, and for $1 \leq \alpha \leq d$ we have $\|H^{(\alpha)}\|_F \leq J$ for some normalised F-function F. Define $T(x)$ to be the local Schrieffer-Wolff transformation as constructed in Section 4.4.2, so that $e^{T(x)}H'(x)e^{-T(x)}$ is block-diagonal with respect to the projector $P_0 := \otimes_{i\in\Gamma_{\mathsf{anc}}} |0_i\rangle\langle 0_i|$. Define the Hamiltonian $H_{\mathsf{eff}}(x)$ on $\mathcal{H}_{\mathsf{eff}} := \otimes_{i\in\Gamma_{\mathsf{eff}}} \mathcal{H}_i$ by

$$H_{\mathsf{eff}}(x) \otimes P_0 := P_0 e^{T(x)} H'(x) e^{-T(x)} P_0 \ . \tag{4.225}$$

If $H_{\mathsf{eff}}(x)$ is analytic in x, then we say that $H'(x)$ is a gadget of degree d for the target Hamiltonian $H_{\mathrm{tar}} := H_{\mathsf{eff}}(0)$.

Note that the condition that $H_{\mathsf{eff}}(x)$ is analytic is equivalent to requiring that, for $\{V^{(q)}\}_{q\geq 1}$ defined via Eq. (4.161), we have $P_0 V^{(q)} P_0 = 0$ for $1 \leq q \leq d-1$. In Appendix 4.6.4, we prove several general results about gadgets as defined in Definition 4.4.7. In particular, we show that the definition is compatible with previous notions of perturbative simulation [87, 56, 1] (Lemma 4.6.11), we show that gadgets can be combined in parallel provided that $d \leq 3$ (Theorem 4.6.13), and we give a strict bound on their locality (Lemma 4.6.15).

A remark on notation and terminology: typically, when $H'(x)$ is an extensive Hamiltonian (i.e. when $\Gamma_{\mathsf{anc}} = \mathcal{O}(n)$) simulating many interactions, it is referred to as a *simulator* rather than a gadget. Usually such $H'(x)$ are built by the parallel application of many gadgets each simulating local terms (as is made precise in Theorem 4.6.13), which are denoted with lowercase $h'_i(x)$. Definition 4.4.7 (and some of our subsequent results) slightly misuse this convention, and we use uppercase H to emphasise that the statements are also valid for extensive simulator Hamiltonians.

Properties of simulator Hamiltonians

For convenience, we collect our main results about the local Schrieffer-Wolff transformation, adapted to simulator Hamiltonians, into a Theorem 4.4.9. We separately state the setup required (essentially, that $H'(x)$ is a simulator Hamiltonian for some H_{tar}) as Assumption 4.4.8, as these conditions will be re-used in Section 4.5.

Assumption 4.4.8. *Let $H'(x)$ be a family of Hamiltonians on the system $\mathcal{H}' = \otimes_{i\in\Gamma_{\mathsf{eff}}\cup\Gamma_{\mathsf{anc}}}\mathcal{H}_i$, of the form*

$$H'(x) = x^{-d}\sum_{\alpha=0}^{d} x^{\alpha} H^{(\alpha)} \ , \tag{4.226}$$

where $H^{(0)} = \Delta\sum_{i\in\Gamma_{\mathsf{anc}}}(\mathbb{1} - |0_i\rangle\langle 0_i|)$ for some states $|0_i\rangle \in \mathcal{H}_i$, and where for $1 \leq \alpha \leq d$ we have $\|H^{(\alpha)}\|_F \leq J$ for some normalised F-function F on $\Gamma = \Gamma_{\mathsf{eff}} \cup \Gamma_{\mathsf{anc}}$. Let $P_0 := \otimes_{i\in\Gamma_{\mathsf{anc}}}|0_i\rangle\langle 0_i|$, and assume that $H'(x)$ is a gadget of degree d for some Hamiltonian H_{tar} on $\mathcal{H}_{\mathsf{eff}} := \otimes_{i\in\Gamma_{\mathsf{eff}}}\mathcal{H}_i$, with associated Schrieffer-Wolff transformation $T(x)$. Assume that $x \leq x_* := 1/(\theta(2+\|F\|))$, where $\theta = 64(1+4J/\Delta)^2$ perturbative series converges and we can define

$$H_{\mathsf{eff}}(x)\otimes P_0 = P_0 e^{T(x)} H'(x) e^{-T(x)} P_0 \ . \tag{4.227}$$

Theorem 4.4.9 (Properties of simulator Hamiltonians)**.** *Let $H'(x)$ be a family of Hamiltonians satisfying the conditions of Assumption 4.4.8. Then, for $x \leq x_*$, the following holds:*

(I) Both $H_{\mathsf{eff}}(x)$ and $T(x)$ are analytic functions of x, and we have the locality bounds

$$\|H_{\mathsf{eff}}(x)\|_F \leq \frac{\Delta\theta^d}{8} \ , \quad \|T(x)\|_F \leq \frac{\theta x}{8} \ . \tag{4.228}$$

(II) Let $\rho'_\beta(x)$ and $\rho_{\beta,\mathsf{eff}}(x)$ be the Gibbs states of $H'(x)$ and $H_{\mathsf{eff}}(x)$ respectively at inverse temperature $\beta > 0$. Then

$$\left\|\rho'_\beta(x) - e^{-T(x)}(\rho_{\beta,\mathsf{eff}}(x) \otimes P_0)e^{T(x)}\right\|_1 \leq 2(m-1)|\Gamma_{\mathsf{anc}}|e^{-\beta\Delta x^{-d}/2} , \qquad (4.229)$$

where $m = \max_{x\in\Gamma_{\mathsf{anc}}} \dim \mathcal{H}_x$.

(III) Let $P'_{\mathsf{ground}}(x)$ and $P_{\mathsf{ground},\mathsf{eff}}(x)$ be the projectors onto the ground spaces of $H'(x)$ and $H_{\mathsf{eff}}(x)$ respectively. Then

$$P'_{\mathsf{ground}}(x) = e^{-T(x)}(P_{\mathsf{ground},\mathsf{eff}}(x) \otimes P_0)e^{T(x)} . \qquad (4.230)$$

Proof of Theorem 4.4.9. For consistency with Section 4.4.2, we write $H(x)$ for the normalisation of $H'(x)$ by a factor of x^d, leading to the non-singular Hamiltonian

$$H(x) := x^d H'(x) = \sum_{\alpha=0}^{d} x^\alpha H^{(\alpha)} . \qquad (4.231)$$

We may then construct the series $\{T^{(q)}\}_{q\geq 1}$, $\{V^{(q)}\}_{q\geq 1}$ as in Eqs. (4.161) and (4.175). For x small enough so that the series $T(x) := \sum_{q\geq 1} x^q T^{(q)}$ and $V(x) := \sum_{q\geq 1} x^q V^{(q)}$ both converge, we have that $H(x)$ (and hence $H'(x)$) is block-diagonal with respect to $e^{-T(x)}P_0e^{T(x)}$, and we write

$$P_0 e^{T(x)} H'(x) e^{-T(x)} P_0 = x^{-d} \sum_{q\geq 1} x^q P_0 V^{(q)} P_0 . \qquad (4.232)$$

By the assumption that $H'(x)$ is a gadget Hamiltonian, we are guaranteed that the right-hand side of this expression is analytic in x and hence $P_0V^{(q)}P_0 = 0$ for $q < d$. Therefore we can write

$$P_0 e^{T(x)} H'(x) e^{-T(x)} P_0 = \sum_{q\geq 0} x^q P_0 V^{(q+d)} P_0 = H_{\mathsf{eff}}(x) \otimes P_0 . \qquad (4.233)$$

Expanding into powers of x, we write

$$H_{\mathsf{eff}}(x) = \sum_{q\geq 0} x^q H^{(q)}_{\mathsf{eff}} , \qquad (4.234)$$

where $H^{(q)}_{\mathsf{eff}} \otimes P_0 = P_0 V^{(q+d)} P_0$.

(I) The analyticity of $H_{\mathsf{eff}}(x)$ and $T(x)$ for $x \leq x_*$ is immediate from their construction as power series in x. In order to bound $\|H_{\mathsf{eff}}(x)\|_F$, we must relate the $\|\cdot\|_F$-norm of $H^{(q)}_{\mathsf{eff}}$ to that of $V^{(q+d)}$. Expanding $V^{(q+d)}$ into a sum of local terms,

$$V^{(q+d)} = \sum_{A\subseteq\Gamma} V^{(q+d)}_A = \sum_{A_{\mathsf{eff}}\subseteq\Gamma_{\mathsf{eff}}} \sum_{A_{\mathsf{anc}}\subseteq\Gamma_{\mathsf{anc}}} V^{(q+d)}_{A_{\mathsf{eff}}\cup A_{\mathsf{anc}}} , \qquad (4.235)$$

we can write

$$H_{\mathsf{eff}}^{(q)} = \sum_{A_{\mathsf{eff}} \subseteq \Gamma_{\mathsf{eff}}} H_{\mathsf{eff},A_{\mathsf{eff}}}^{(q)} \ , \tag{4.236}$$

where

$$H_{\mathsf{eff},A_{\mathsf{eff}}}^{(q)} := \sum_{A_{\mathsf{anc}} \subseteq \Gamma_{\mathsf{anc}}} (\mathbb{1} \otimes \langle \mathbf{0}_{A_{\mathsf{anc}}}|) V_{A_{\mathsf{eff}} \cup A_{\mathsf{anc}}}^{(q+d)} (\mathbb{1} \otimes |\mathbf{0}_{A_{\mathsf{anc}}}\rangle) \ , \tag{4.237}$$

and where $|\mathbf{0}_{A_{\mathsf{anc}}}\rangle = \otimes_{i \in A_{\mathsf{anc}}} |0_i\rangle$ is the unique state on $\bigotimes_{i \in \Gamma_{\mathsf{anc}}} \mathcal{H}_i$ in the image of $P_{A_{\mathsf{anc}}} := \otimes_{i \in A_{\mathsf{anc}}} |0_i\rangle\langle 0_i|$. In particular, by the triangle inequality this implies that

$$\|H_{\mathsf{eff},A_{\mathsf{eff}}}^{(q)}\| \leq \sum_{A_{\mathsf{anc}} \subseteq \Gamma_{\mathsf{anc}}} \|V_{A_{\mathsf{eff}} \cup A_{\mathsf{anc}}}^{(q+d)}\| \ . \tag{4.238}$$

Hence

$$\|H_{\mathsf{eff}}^{(q)}\|_F = \sup_{i,j \in \Gamma_{\mathsf{eff}}} \frac{1}{F(\mathrm{dist}(i,j))} \sum_{\substack{A_{\mathsf{eff}} \subseteq \Gamma_{\mathsf{eff}} \\ i,j \in A_{\mathsf{eff}}}} \|H_{\mathsf{eff},A_{\mathsf{eff}}}^{(q)}\| \tag{4.239}$$

$$\leq \sup_{i,j \in \Gamma_{\mathsf{eff}}} \frac{1}{F(\mathrm{dist}(i,j))} \sum_{\substack{A_{\mathsf{eff}} \subseteq \Gamma_{\mathsf{eff}} \\ i,j \in A_{\mathsf{eff}}}} \sum_{A_{\mathsf{anc}} \subseteq \Gamma_{\mathsf{anc}}} \|V_{A_{\mathsf{eff}} \cup A_{\mathsf{anc}}}^{(q+d)}\| \tag{4.240}$$

$$\leq \sup_{i,j \in \Gamma_{\mathsf{eff}} \cup \Gamma_{\mathsf{anc}}} \frac{1}{F(\mathrm{dist}(i,j))} \sum_{\substack{A \subseteq \Gamma_{\mathsf{eff}} \cup \Gamma_{\mathsf{anc}} \\ i,j \in A}} \|V_A^{(q+d)}\| \tag{4.241}$$

$$= \|V^{(q+d)}\|_F \ . \tag{4.242}$$

By Lemma 4.4.2, we can bound $\|V^{(q)}\|_F \leq \Delta\theta^q/16$. Hence, since $x \leq 1/(2\theta) \leq x_*$,

$$\begin{aligned} \|H_{\mathsf{eff}}(x)\|_F &\leq \sum_{q \geq 0} x^q \|H_{\mathsf{eff}}^{(q)}\|_F \\ &\leq \sum_{q \geq 0} x^q \|V^{(q+d)}\|_F \\ &\leq \frac{\Delta\theta^d}{16} \sum_{q \geq 0} (x\theta)^q \\ &\leq \frac{\Delta\theta^d}{8} \ , \end{aligned}$$

as required. The bound $\|T(x)\|_F \leq \theta x/8$ is immediate from Corollary 4.4.3.

(II) This follows by applying Corollary 4.4.5 to the Hamiltonian $H(x)$, and absorbing the factor of x^{-d} into the inverse temperature, yielding

$$\left\| \rho'_\beta(x) - e^{-T(x)} (\rho_{\beta_{\mathsf{eff}}(x)} \otimes P_0) e^{T(x)} \right\|_1 \leq 2(m-1)|\Gamma_{\mathsf{anc}}| \exp\left(-\beta\Delta x^{-d} (1 - 3\theta\|F\|x/8) \right) \ . \tag{4.243}$$

Since we assume that $x \leq x_* \leq 1/(\theta\|F\|) \leq 4/(3\theta\|F\|)$, this furhter simplifies to

$$\left\|\rho'_\beta(x) - e^{-T(x)}(\rho_{\beta_{\mathsf{eff}}(x)} \otimes P_0)e^{T(x)}\right\|_1 \leq 2(m-1)|\Gamma_{\mathsf{anc}}|e^{-\beta\Delta x^{-d}/2} \ , \qquad (4.244)$$

as required.

(III) This is an immediate consequence of Corollary 4.4.6.

■

4.5 Extrapolation with simulator Hamiltonians

4.5.1 Main statement

In this section, we combine the local Schrieffer-Wolff perturbation theory tools from Section 4.4 with the extrapolation results from Section 4.3. The main result, Theorem 4.5.1 below, establishes analytic approximations for observables on simulator Hamiltonians. In Corollary 4.5.2, we explicitly rephrase this in terms of the simulation overhead required for Richardson extrapolation.

Theorem 4.5.1. *Let $H'(x)$ be a family of degree-d gadget Hamiltonians on $\mathcal{H}' = \bigotimes_{i\in\Gamma_{\mathsf{eff}}\cup\Gamma_{\mathsf{anc}}}\mathcal{H}_i$ satisfying the conditions of Assumption 4.4.8 with exponentially decaying interactions (i.e. with $\|\cdot\|_{F_g}$ in place of $\|\cdot\|_F$, for linear g). Define the target Hamiltonian $H_{\mathrm{tar}} := H_{\mathsf{eff}}(0)$, and let O_A be an observable supported on $A \subseteq \Gamma_{\mathsf{eff}}$, where $|A|, \|O_A\| = \mathcal{O}(1)$. Then the following holds:*

(I) Let $\beta > 0$, and let $\rho'_\beta(x)$, $\rho_{\beta,\mathsf{eff}}(x)$ and $\rho_{\beta,\mathrm{tar}}$ be the Gibbs states corresponding to the Hamiltonians $H'(x)$, $H_{\mathsf{eff}}(x)$, and H_{tar} respectively. Assume that the family $\rho_{\beta,\mathsf{eff}}(x)$ satisfies the exponential correlation decay of Assumption 4.3.1(I) for $x \in [0, x_*]$. Then the function

$$f'_\beta(x) := \begin{cases} \mathrm{tr}[O_A\rho'_\beta(x)] & \text{for} \quad x \in (0, x_*) \\ \mathrm{tr}[O_A\rho_{\beta,\mathrm{tar}}] & \text{for} \quad x = 0 \end{cases} \ , \qquad (4.245)$$

has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = \exp\left(\mathcal{O}(\log^D(\delta^{-1}))\right) \ , \quad R = \min\left\{1/\,\mathcal{O}(\log^D(\delta^{-1})), 1/\,\mathcal{O}(\log^{1/d}(n\delta^{-1}))\right\} \ . \qquad (4.246)$$

(II) Assume that $H_{\mathsf{eff}}(x)$ has a spectral gap as in Assumption 4.3.1(II) for $x \in [0, x_*]$, and let $|\psi'_0(x)\rangle$, $|\psi_{0,\mathsf{eff}}(x)\rangle$ and $|\psi_{0,\mathrm{tar}}\rangle$ be the ground states corresponding to the Hamiltonians $H'(x)$, $H_{\mathsf{eff}}(x)$, and H_{tar} respectively[2]. Then the function

$$f'_{\mathsf{ground}}(x) := \begin{cases} \langle\psi'_0(x)|O_A|\psi'_0(x)\rangle & \text{for} \quad x \in (0, x_*) \\ \langle\psi_{0,\mathrm{tar}}|O_A|\psi_{0,\mathrm{tar}}\rangle & \text{for} \quad x = 0 \end{cases} \ , \qquad (4.247)$$

has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = \exp\left(\mathcal{O}(\log^D(\delta^{-1}))\right) \ , \quad R = 1/\,\mathcal{O}(\log^{D+1}(\delta^{-1})) \ , \qquad (4.248)$$

[2]Note that the uniqueness of the ground state for $H'(x)$ follows immediately from our assumption on $H_{\mathsf{eff}}(x)$ and Theorem 4.4.9(III).

(Proof on page 150.)

Note that the ground state part is n-independent, as in Theorem 4.3.2, however for Gibbs states we obtain a $1/\mathcal{O}(\log^{1/d}(n\Delta^{-1}))$ upper bound on R — this is necessary to ensure that the Gibbs state lies in the correct effective subspace (which does not require n-dependent scaling for ground states due to Corollary 4.4.6).

Corollary 4.5.2 (Extrapolating local observables with simulator Hamiltonians)**.** *The value of* $\mathrm{tr}[O_A\rho_{\beta,\mathrm{tar}}]$ *(respectively* $\langle\psi_{0,\mathrm{tar}}|O_A|\psi_{0,\mathrm{tar}}\rangle$*) can be calculated up to any desired accuracy* $\epsilon > 0$*, given the values of* $f'_\beta(x_k)$ *(respectively* $f'_{\mathsf{ground}}(x_k)$*) at* m *Chebyshev sample points* $\{x_k\}_{k=1}^m$*, where each* x_k *is bounded above zero by* $x_{\min} := \min_k x_k$*, where:*

(I) For f'_β*,*

$$m = \mathcal{O}(\log^D(\epsilon^{-1})) \ , \quad x_{\min} = \frac{\min\{1/\log^D(\epsilon^{-1}), 1/\log^{1/d}(n\epsilon^{-1})\}}{\log^{2D}(\epsilon^{-1})} \ . \tag{4.249}$$

(II) For f'_{ground}*,*

$$m = \mathcal{O}(\log^D(\epsilon^{-1})) \ , \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{3D+1}(\epsilon^{-1}))} \ . \tag{4.250}$$

In particular, this process only requires simulator Hamiltonians with interaction strengths scaling as $\sim x_{\min}^{-d}$*, which is* $\mathrm{poly}\log(n\epsilon^{-1})$ *for* f'_β *and* $\mathrm{poly}(\log(\epsilon^{-1}))$ *for* f'_{ground}*. These conclusions also hold using noisy estimates of* $f'_\beta(x_k)$ *(respectively* $f'_{\mathsf{ground}}(x_k)$*) each with additive error* $\delta = \Theta(\epsilon/\log\log(\epsilon^{-1}))$*.*

Proof of Corollary 4.5.2. This follows from Theorem 4.5.1 and Corollary 4.2.4, via a similar argument to the proof of Corollary 4.3.3. Concretely, we aim to choose parameters M, m, and δ, such that

$$\epsilon = (\delta + 2^{-m}M)\,\mathcal{O}(\log m) \ , \tag{4.251}$$

where Theorem 4.5.1 fixes $M = \exp(\mathcal{O}(\log^D(\delta^{-1}))$. This is achievable by taking $\delta = \epsilon/\mathcal{O}(\log\log(\epsilon^{-1}))$ and $m = \mathcal{O}(\log^D(\epsilon^{-1}))$. Using the relationship $x_{\min} \sim R/m^2$ with the values of R given by Theorem 4.5.1 gives the required result. ∎

Our proof of Theorem 4.5.1 will rely on the observation that Theorem 4.4.9(II)-(III) ensures that

$$f'_\beta(x) \approx \mathrm{tr}[O_{A,\mathsf{eff}}(x)\rho_{\beta,\mathsf{eff}}(x)] \quad \text{and} \quad f'_{\mathsf{ground}}(x) = \langle\psi_{0,\mathsf{eff}}(x)|O_{A,\mathsf{eff}}(x)|\psi_{0,\mathsf{eff}}(x)\rangle \ , \tag{4.252}$$

where $O_{A,\mathsf{eff}}(x)$ is defined as

$$O_{A,\mathsf{eff}}(x) := (\mathbb{1}_{\mathsf{eff}} \otimes \langle\mathbf{0}_{\mathsf{anc}}|)e^{T(x)}O_A e^{-T(x)}(\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle) \ , \tag{4.253}$$

where $|\mathbf{0}_{\mathsf{anc}}\rangle$ is the unique state on $\mathcal{H}_{\mathsf{anc}}$ such that $P_0 = |\mathbf{0}_{\mathsf{anc}}\rangle\langle\mathbf{0}_{\mathsf{anc}}|$. In this way, we can reduce the extrapolation task in Theorem 4.5.1 — which involves the singular Hamiltonian $H'(x)$ — to an extrapolation task on the well-behaved (analytic) Hamiltonian $H_{\mathsf{eff}}(x)$. From here, the proof follows exactly as in Theorems 4.3.9 and 4.3.14. The only additional subtlety is that we must account for the fact that $O_{A,\mathsf{eff}}(x)$ no longer has constant-sized support, but we can use the locality of $T(x)$ given by Theorem 4.4.9(I) with Lieb-Robinson bounds to argue that its support is approximately localised around A, made precise in the following lemma, which is proved in Appendix 4.6.4:

4.5.2 Quasi-locality of effective observables

Lemma 4.5.3 (See Lemma 4.6.16)**.** *Let $O_{A,\mathsf{eff}}(x)$ be the observable on Γ_{eff} defined in Eq.* (4.253)*, and assume $x \leq 1/(2\theta)$. Then, for any $r \geq 0$, there exists an observable $O^{[r]}_{A,\mathsf{eff}}(x)$ on Γ_{eff}, such that $O^{[r]}_{A,\mathsf{eff}}(x)$ is supported on $B_r(A)$, we have*

$$\left\| O_{A,\mathsf{eff}}(x) - O^{[r]}_{A,\mathsf{eff}}(x) \right\| \leq a e^{-bg(r)} \ , \tag{4.254}$$

for some constants $a, b > 0$. Moreover, $O^{[r]}_{A,\mathsf{eff}}(x)$ extends to a complex function $O^{[r]}_{A,\mathsf{eff}}(z)$ which is analytic on the disc $|z| \leq 1/(2\theta)$, and which is bounded as

$$\sup_{|z|\leq 1/(2\theta)} \|O^{[r]}_{A,\mathsf{eff}}(z)\| \leq a' e^{b' r^D} \ , \tag{4.255}$$

for constants $a', b' > 0$.

In fact, Lemma 4.5.3 is sufficient to guarantee that $O_{A,\mathsf{eff}}(x)$ remains (approximately) local even when transformed by the quantum belief propagation and spectral flow operators. This is stated in the following lemma, which can be viewed as a generalisation of Lemma 4.2.15, and which is proved in Appendix 4.6.5.

Lemma 4.5.4 (See Lemma 4.6.19)**.** *Let H be a Hamiltonian with bounded $\|\cdot\|_{F_g}$-norm for g linear, and let O be an observable with $\|O\| = \mathcal{O}(1)$ localised around $i \in \Gamma$ in the following sense: for every $r \geq 0$, there exists an observable $O^{[r]}$ with support contained within $B_r(\{i\})$ such that*

$$\|O - O^{[r]}\| \leq a e^{-bg(r)} \ , \tag{4.256}$$

for some constants $a, b > 0$. Let $\Phi_H(O)$ and $\Psi_H(O)$ be the quantum belief propagation and spectral flow operators as defined in Eqs. (4.29) *and* (4.35)*. Then there exist constants $a'_1, a'_2, b'_1, b'_2 > 0$ such that, for every $r \geq 0$, there exist operators $\Phi^{[r]}_H(O)$ and $\Psi^{[r]}_H(O)$ with support contained within $B_r(\{i\})$ such that*

$$\|\Phi_H(O) - \Phi^{[r]}_H(O)\| \leq a'_1 \|O\| e^{-b'_1 g(r)} \ , \tag{4.257}$$

$$\|\Psi_H(O) - \Psi^{[r]}_H(O)\| \leq a'_2 \|O\| e^{-b'_2 g(r)/\log^2 g(r)} \ . \tag{4.258}$$

We are now ready to prove Theorem 4.5.1 below.

Proof of Theorem 4.5.1. We define the functions $f_{\beta,\mathsf{eff}}(x)$ and $f_{\mathsf{ground},\mathsf{eff}}(x)$ for $x \in [0, x_*]$ by

$$f_{\beta,\mathsf{eff}}(x) := \mathrm{tr}[O_{A,\mathsf{eff}}(x)\rho_{\beta,\mathsf{eff}}(x)] \tag{4.259}$$

$$f_{\mathsf{ground},\mathsf{eff}}(x) := \langle\psi_{0,\mathsf{eff}}(x)|O_{A,\mathsf{eff}}(x)|\psi_{0,\mathsf{eff}}(x)\rangle \ , \tag{4.260}$$

where $O_{A,\mathsf{eff}}(x)$ is defined as in Eq. (4.253). By Theorem 4.4.9(II), we can bound

$$|f'_\beta(x) - f_{\beta,\mathsf{eff}}(x)| = \mathrm{tr}\left[O_A\left(\rho'_\beta(x) - e^{-T(x)}(\rho_{\beta,\mathsf{eff}}(x)\otimes P_0)e^{T(x)}\right)\right] \tag{4.261}$$

$$\leq \|O_A\| \left\|\rho'_\beta(x) - e^{-T(x)}(\rho_{\beta,\mathsf{eff}}(x)\otimes P_0)e^{T(x)}\right\|_1 \tag{4.262}$$

$$\leq 2(m-1)\|O_A\||\Gamma_{\mathsf{anc}}|e^{-\beta\Delta x^{-d}/2} \ , \tag{4.263}$$

whilst $f'_{\mathsf{ground}}(x) = f_{\mathsf{ground},\mathsf{eff}}(x)$ is guaranteed by Theorem 4.4.9(II). Our problem thus reduces to finding analytic approximations for $f_{\beta,\mathsf{eff}}(x)$ and $f_{\mathsf{ground},\mathsf{eff}}(x)$. Such analytic approximations exist by Theorem 4.3.2, however there are a couple of additional subtleties in this case:

- Firstly, Theorem 4.3.2 is stated in terms of strictly local O_A supported on $|A| = \mathcal{O}(1)$ sites. This is not satisfied by $O_{A,\mathsf{eff}}(x)$, however Lemma 4.5.3 ensures that $O_{A,\mathsf{eff}}(x)$ can be truncated at finite radius up to exponentially small errors, and thus Lemma 4.5.4 ensures that the corresponding quantum belief propagation and spectral flow operators can be truncated with the same asymptotic behaviour as for strictly local O_A, as given by Lemma 4.2.15. Since the strict locality of O_A is only necessary in order to obtain these truncation bounds, the proof of Theorem 4.3.2 goes through unchanged, except using Lemma 4.5.4 in place of Lemma 4.2.15.

- Secondly, in the statement of Theorem 4.3.2, it is assumed that $\|O_A\| = \mathcal{O}(1)$, and this results in $M = \mathcal{O}(\|O_A\|) = \mathcal{O}(1)$ for both the Gibbs state and ground state parts. In our case, $\|O_{A,\mathsf{eff}}(x)\|$ may grow for complex $x \in \mathbb{C}$, however the truncated observables $\|O^{[r]}_{A,\mathsf{eff}}(x)\|$ can be bounded following Lemma 4.5.3 as $\sim e^{\mathcal{O}(r^D)}$. We ultimately take $r = \mathcal{O}(\log \delta^{-1})$ in the proof of Theorem 4.3.2, and hence an additional factor of $\exp(\mathcal{O}(\log^D(\delta^{-1})))$ will appear in the M expressions.

We deal with the quantitative conclusions of this for $f_{\beta,\mathsf{eff}}(x)$ and $f_{\mathsf{ground},\mathsf{eff}}(x)$ separately:

(I) For $\delta > 0$, Theorem 4.3.2(I) gives a $(\delta/2, M, R')$-analytic approximation $\tilde{f}_{\beta,\mathsf{eff}}(z)$ for $f_{\beta,\mathsf{eff}}(x)$, where

$$M = \exp\left(\mathcal{O}(\log^D(\delta^{-1}))\right) \ , \quad R' = 1/\,\mathcal{O}(\log^D(\delta^{-1})) \ . \tag{4.264}$$

As mentioned, the $e^{\mathcal{O}(\log^D(\delta^{-1}))}$ factor in M arises from the corresponding bound on $\|O^{[r]}_{A,\mathsf{eff}}(z)\|$ given by Eq. (4.255), as we ultimately take $r \sim \log(\delta^{-1})$ in the

proof of Theorem 4.3.2. If we take $x \sim 1/\log^{1/d}(n\delta^{-1})$, the right-hand side of Eq. (4.263) can be upper bounded by $\delta/2$, implying that $\tilde{f}_{\beta,\mathsf{eff}}(z)$ is a (δ, M, R)-analytic approximation for $f'_\beta(x)$, where

$$M = \exp\left(\mathcal{O}(\log^D(\delta^{-1}))\right) \ , \quad R = \min\left\{1/\mathcal{O}(\log^D(\delta^{-1})), 1/\mathcal{O}(\log^{1/d}(n\delta^{-1}))\right\} \ . \tag{4.265}$$

(II) For $\delta > 0$, Theorem 4.3.2(II) gives a (δ, M, R)-analytic approximation $\tilde{f}_{\mathsf{ground},\mathsf{eff}}(z)$ for $f_{\mathsf{ground},\mathsf{eff}}(x)$ (and hence $f'_{\mathsf{ground}}(x)$), where

$$M = \exp\left(\mathcal{O}(\log^D(\delta^{-1}))\right) \ , \quad R = 1/\mathcal{O}(\log^{D+1}(\delta^{-1})) \ , \tag{4.266}$$

by similar arguments to the previous case.

■

4.5.3 Example: locality reduction

As a concrete example of the types of simulations made possible through the sorts of perturbative gadgets described by this work, we mention here the specific application of reducing 3-local Hamiltonians to 2-local Hamiltonians [57] (more generally, k-local to 2-local reduction is possible for any $k \geq 3$ [72] by recursive simulation). Such reductions are made possible through the use of a 3-to-2 gadget, for example as introduced in Ref. [72]. This gives a degree-3 gadget $h'(x)$ on $\mathcal{H}_{\mathsf{eff}} \otimes \mathcal{H}_{\mathsf{anc}}$, where $\mathcal{H}_{\mathsf{eff}} \simeq (\mathbb{C}^2)^{\otimes 3}$, $\mathcal{H}_{\mathsf{anc}} \simeq \mathbb{C}^2$, which simulates an arbitrary interaction of the form $h := h_1 \otimes h_2 \otimes h_3$ on $\mathcal{H}_{\mathsf{eff}}$, whilst $h'(x)$ is itself only 2-local. Explicitly, $h'(x)$ is defined as

$$\begin{aligned} h'(x) = \ & x^{-3}\, \mathbb{1} \otimes \mathbb{1} \otimes \mathbb{1} \otimes |1\rangle\langle 1| \\ & + x^{-2} \frac{1}{\sqrt{2}} \left(-h_1 \otimes \mathbb{1} + \mathbb{1} \otimes h_2\right) \otimes \mathbb{1} \otimes X - \mathbb{1} \otimes \mathbb{1} \otimes h_3 \otimes |1\rangle\langle 1| \\ & + x^{-1} \frac{1}{2} \left(-h_1 \otimes \mathbb{1} + \mathbb{1} \otimes h_2\right)^2 \otimes \mathbb{1} \otimes \mathbb{1} \\ & + \frac{1}{2} \left(h_1^2 \otimes \mathbb{1} + \mathbb{1} \otimes h_2^2\right) \otimes h_3 \otimes \mathbb{1} \ . \end{aligned} \tag{4.267}$$

The fact that $h'(x)$ is a gadget for h follows from direct calculation of the local Schrieffer-Wolff perturbative series (or otherwise can be deduced by the previous analysis of Ref. [72] along with our result Lemma 4.6.11). Since $h'(x)$ is a gadget of degree $d \leq 3$, many such gadgets can be combined in parallel (by our gadget combination result Lemma 4.6.13) to simulate an arbitrary 3-local Hamiltonian H_{tar} with a 2-local simulator Hamiltonian $H'(x)$. According to Corollary 4.5.2, the local properties of the ground and Gibbs states of H can be extrapolated from the analogous properties of $H'(x)$ for $x \sim 1/\operatorname{poly}\log(\epsilon^{-1})$, compared to the $1/\operatorname{poly}(n, \epsilon^{-1})$ scalings necessary for exact full-spectrum simulation.

The important caveat for this comparison is that the quasi-local effective Hamiltonian $H_{\mathsf{eff}}(x)$ obtained from $H'(x)$ via the local Schrieffer-Wolff transformation must remain uniformly non-critical along the path of extrapolation, as described in Assumption 4.3.1. This is in general difficult to prove rigorously, but we may heuristically expect this to hold for many disordered systems (see e.g. Ref. [216]); the numerics of Ref. [198] show that their similar techniques work well empirically in many cases when non-criticality cannot be rigorously proved.

Acknowledgements

We acknowledge financial support from the Novo Nordisk Foundation (Grant No. NNF20OC0059939 'Quantum for Life' and Grant No. NNF25OC0105181 'Molecular Recognition from Quantum Computing'), and VILLUM FONDEN via the QMATH Centre of Excellence (Grant No.10059).

4.6 Appendix

4.6.1 Bounding the condition number

Below we give an elementary proof of the claim from Section 4.2.1, that the condition number $\alpha(\mathbf{x})$ arising from Chebyshev nodes $\{x_k\}_{k=1}^m$ is bounded as $\mathcal{O}(\log(m))$.

Lemma 4.6.1. *Let $m \geq 2$, and define $\{x_k\}_{k=1}^m$ by*

$$x_k = \sin^2\left(\frac{(2k-1)\pi}{4m}\right) \ . \tag{4.268}$$

and let

$$\alpha(\mathbf{x}) := \sum_{k=1}^{m} \prod_{j\neq k} \left|\frac{x_j}{x_j - x_k}\right| \ . \tag{4.269}$$

Then we can bound

$$\alpha(\mathbf{x}) \leq 3\log(m) \ . \tag{4.270}$$

Proof of Lemma 4.6.1. Using the identity $\sin^2\theta - \sin^2\phi = \sin(\theta+\phi)\sin(\theta-\phi)$, we have

$$x_j - x_k = \sin\left(\frac{(2j-1)\pi + (2k-1)\pi}{4m}\right)\sin\left(\frac{(2j-1)\pi - (2k-1)\pi}{4m}\right) \tag{4.271}$$

$$= \sin\left(\frac{(j+k-1)\pi}{2m}\right)\sin\left(\frac{(j-k)\pi}{2m}\right) \ , \tag{4.272}$$

and hence

$$\prod_{j\neq k}\left|\frac{x_j}{x_j-x_k}\right| = \left|\frac{\left(\prod_{j\neq k}\sin\left(\frac{(2j-1)\pi}{4m}\right)\right)^2}{\left(\prod_{j\neq k}\sin\left(\frac{(j+k-1)\pi}{2m}\right)\right)\left(\prod_{j\neq k}\sin\left(\frac{(j-k)\pi}{2m}\right)\right)}\right| . \tag{4.273}$$

For the evaluation of these products we will make repeated use of the following identity:

$$\prod_{k=1}^{m-1}\sin\left(\frac{k\pi}{m}\right) = \frac{m}{2^{m-1}} . \tag{4.274}$$

For the product in the numerator we have

$$\left(\prod_{j\neq k}\sin\left(\frac{(2j-1)\pi}{4m}\right)\right)^2 = \frac{\prod_{j=1}^{2m}\sin\left(\frac{(2j-1)\pi}{4m}\right)}{\sin^2\left(\frac{(2k-1)\pi}{4m}\right)} \tag{4.275}$$

$$= \frac{\prod_{j=1}^{4m-1}\sin\left(\frac{j\pi}{4m}\right)}{\prod_{j=1}^{2m-1}\sin\left(\frac{j\pi}{2m}\right)}\cdot\frac{1}{\sin^2\left(\frac{(2k-1)\pi}{4m}\right)} \tag{4.276}$$

$$= \frac{1}{2^{2m-1}\sin^2\left(\frac{(2k-1)\pi}{4m}\right)} . \tag{4.277}$$

In the first line we used that the terms corresponding to j and $m-j$ are identical, and in the third line we used Eq. (4.274) twice.

Meanwhile, for the denominator we have

$$\left|\left(\prod_{j\neq k}\sin\left(\frac{(j+k-1)\pi}{2m}\right)\right)\left(\prod_{j\neq k}\sin\left(\frac{(j-k)\pi}{2m}\right)\right)\right| = \left|\frac{1}{\sin\left(\frac{(2k-1)\pi}{2m}\right)}\prod_{\substack{j=-m\\ j\neq k}}^{m-1}\sin\left(\frac{(j+k)\pi}{2m}\right)\right| \tag{4.278}$$

$$= \frac{1}{\sin\left(\frac{(2k-1)\pi}{2m}\right)}\prod_{j=1}^{2m-1}\sin\left(\frac{j\pi}{2m}\right) \tag{4.279}$$

$$= \frac{2m}{2^{2m-1}\sin\left(\frac{(2k-1)\pi}{2m}\right)} , \tag{4.280}$$

where in the last line we used Eq. (4.274). Hence

$$\prod_{j\neq k}\left|\frac{x_j}{x_j-x_k}\right| = \frac{\sin\left(\frac{(2k-1)\pi}{2m}\right)}{2m\sin^2\left(\frac{(2k-1)\pi}{4m}\right)} \tag{4.281}$$

$$= \frac{1}{m}\cot\left(\frac{(2k-1)\pi}{4m}\right) . \tag{4.282}$$

We can then write

$$\alpha(\mathbf{x}) = \frac{1}{m}\sum_{k=1}^{m} \cot\left(\frac{(2k-1)\pi}{4m}\right) \ . \tag{4.283}$$

Using that $\cot\theta \leq \theta^{-1}$ for $\theta \in (0, \pi/2)$, we then have

$$\alpha(\mathbf{x}) \leq \frac{4}{\pi}\sum_{k=1}^{m} \frac{1}{2k-1} \tag{4.284}$$

$$\leq 3\log(m) \ , \tag{4.285}$$

where the last inequality holds for $m \geq 2$. ∎

4.6.2 Localising quantum belief propagation and the spectral flow

In this section, we will prove some basic results involving local truncations of the quantum belief propagation and spectral flow operators defined in Section 4.2.2. The qualitative conclusions of this section are not novel and are stated and proved here for completeness and consistency with our conventions; the locality of the quantum belief propagation operator has been informally proved as early as Ref. [211], Corollary 3, and more recently applied to prove the GALI property for Gibbs states in Ref. [193]. Meanwhile, a similar result concerning the locality of the spectral flow operator appears in Ref. [199] Section 6.5, which this work follows closely.

Alternative representation of the spectral flow

Firstly, we prove the following Lemma, stated in the main text as Lemma 4.2.10, which expresses the spectral flow operator $\Psi_H(X)$ in a form which will be helpful for our subsequent analysis.

Lemma 4.6.2 (Restatement of Lemma 4.2.10)**.** *The spectral flow operator $\Psi_H(O)$ as defined in Lemma 4.2.9 can be written as*

$$\Psi_H(X) = \int_{-\infty}^{\infty} \tilde{w}_\gamma(t) e^{itH} X e^{-itH} \ , \tag{4.286}$$

where $\tilde{w}_\gamma : \mathbb{R} \to \mathbb{R}$ is an odd L_1 function with $|\tilde{w}_\gamma(t)| \leq 1/2$ for all t. Moreover, for $|t| \geq e^3\gamma^{-1}$,

$$|\tilde{w}_\gamma(t)| \leq W_1 \left(\frac{\gamma t}{\log^2(\gamma t)}\right)^2 \exp\left(-\frac{2\gamma t}{7\log^2(\gamma t)}\right) \ , \tag{4.287}$$

$$\int_t^{\infty} \mathrm{d}s \tilde{w}_\gamma(s) \leq W_2\gamma^{-1} \left(\frac{\gamma t}{\log^2(\gamma t)}\right)^3 \exp\left(-\frac{2\gamma t}{7\log^2(\gamma t)}\right) \ , \tag{4.288}$$

for some constants $W_1, W_2 > 0$. Moreover,

$$\int_{-\infty}^{\infty} \mathrm{d}t |\tilde{w}_\gamma(t)| \leq W_3\gamma^{-1} \ , \tag{4.289}$$

for some constant $W_3 > 0$.

Proof of Lemma 4.6.2. By choosing the function w_γ given by [205] (which in particular is even and non-negative) and reordering the integrals, we rewrite Eq. (4.35) to give

$$\Psi_H(X) = \int_{-\infty}^{\infty} \mathrm{d}t w_\gamma(t) \int_0^t \mathrm{d}u e^{iuH} X e^{-iuH} \tag{4.290}$$

$$= \int_0^{\infty} \mathrm{d}t w_\gamma(t) \int_0^t \mathrm{d}u e^{iuH} X e^{-iuH} - \int_{-\infty}^{0} \mathrm{d}t w_\gamma(t) \int_t^0 \mathrm{d}u e^{iuH} X e^{-iuH} \tag{4.291}$$

$$= \int_0^{\infty} \mathrm{d}u \int_u^{\infty} \mathrm{d}t w_\gamma(t) e^{iuH} X e^{-iuH} - \int_{-\infty}^{0} \mathrm{d}u \int_{-\infty}^{u} \mathrm{d}t w_\gamma(t) e^{iuH} X e^{-iuH} \tag{4.292}$$

$$= \int_{-\infty}^{\infty} \mathrm{d}t \left(\mathsf{sign}(t) \int_{|t|}^{\infty} \mathrm{d}u w_\gamma(u) \right) e^{itH} X e^{-itH} \ , \tag{4.293}$$

where in the last line we have used that w_γ is an even function, and relabelled $t \leftrightarrow u$. It remains to prove the stated bounds on the function

$$\tilde{w}_\gamma(t) := \int_{|t|}^{\infty} \mathrm{d}u w_\gamma(u) \ . \tag{4.294}$$

The fact that $|\tilde{w}_\gamma(t)| \leq 1/2$ for all t follows by the normalisation of $w_\gamma(t)$.

Using the bound given by Eq. (4.36) for $t \geq e^{-1/\sqrt{2}}\gamma^{-1}$, we have

$$\tilde{w}_\gamma(t) \leq 2(e\gamma)^2 \int_t^{\infty} \mathrm{d}u \cdot u \exp\left(-\frac{2\gamma u}{7 \log^2(\gamma u)} \right) \tag{4.295}$$

$$= 2e^2 \int_t^{\infty} \mathrm{d}u \left(\frac{\log^7(\gamma u)}{\gamma u (\log(\gamma u) - 2)} \right) \left(\gamma^3 u^2 \cdot \frac{\log(\gamma u) - 2}{\log^7(\gamma u)} \right) \exp\left(-\frac{2\gamma u}{7 \log^2(\gamma u)} \right) \ . \tag{4.296}$$

We may bound

$$\frac{\log^7(x)}{x(\log(x) - 2)} \leq 200 \quad \text{for } x \geq e^3 \ , \tag{4.297}$$

and hence assuming that $t \geq e^3 \gamma^{-1}$ we have

$$\tilde{w}_\gamma(t) \leq 400 e^2 \int_t^{\infty} \mathrm{d}u \left(\gamma^3 u^2 \cdot \frac{\log(\gamma u) - 2}{\log^7(\gamma u)} \right) \exp\left(-\frac{2\gamma u}{7 \log^2(\gamma u)} \right) \tag{4.298}$$

$$= 400 e^2 \int_{\frac{\gamma t}{\log^2(\gamma t)}}^{\infty} \mathrm{d}v \cdot v^2 e^{-\frac{2v}{7}} \ , \tag{4.299}$$

where in the second line we performed the change of variable

$$v = \frac{\gamma u}{\log^2(\gamma u)} \Rightarrow \mathrm{d}v = \frac{\gamma(\log(\gamma u) - 2)}{\log^3(\gamma u)} \mathrm{d}u \ . \tag{4.300}$$

Evaluating the integral, this gives

$$\tilde{w}_\gamma(t) \leq 700e^2 \left[e^{-2v/7}(2v^2+14v+49)\right]_{v=\frac{\gamma t}{\log^2(\gamma t)}} \tag{4.301}$$

$$\leq 2100e^2 \left[e^{-2v/7}v^2\right]_{v=\frac{\gamma t}{\log^2(\gamma t)}} \tag{4.302}$$

$$= W_1 \left(\frac{\gamma t}{\log^2(\gamma t)}\right)^2 \exp\left(\frac{7\gamma t}{\log^2(\gamma t)}\right) \ , \tag{4.303}$$

where the second line holds for $\gamma t \geq e^3$, and $W_1 = 2100e^2 \approx 1.55 \times 10^4$.

Using this result, we can bound the integral

$$\int_t^\infty \mathrm{d}s \tilde{w}_\gamma(s) \leq 2100e^2 \int_t^\infty \mathrm{d}s \left(\frac{\gamma s}{\log^2(\gamma s)}\right)^2 \exp\left(-\frac{2\gamma s}{7\log^2(\gamma s)}\right) \tag{4.304}$$

$$= 2100e^2 \int_t^\infty \mathrm{d}s \left(\frac{\gamma^4 s^3(\log(\gamma s)-2)}{\log^9(\gamma s)}\right)\left(\frac{\log^5(\gamma s)}{\gamma^2 s(\log(\gamma s)-2)}\right)\exp\left(-\frac{2\gamma s}{7\log^2(\gamma s)}\right) \ . \tag{4.305}$$

This time, we may bound

$$\frac{\log^5(x)}{x(\log(x)-2)} \leq 25 \quad \text{for } x \geq e^3 \ , \tag{4.306}$$

and hence

$$\int_t^\infty \mathrm{d}s \tilde{w}_\gamma(s) \leq 52500e^2\gamma^{-1} \int_t^\infty \mathrm{d}s \left(\frac{\gamma^4 s^3(\log(\gamma s)-2)}{\log^9(\gamma s)}\right)\exp\left(-\frac{2\gamma s}{7\log^2(\gamma s)}\right) \tag{4.307}$$

$$= 52500e^2\gamma^{-1} \int_{\frac{\gamma t}{\log^2(\gamma t)}}^\infty \mathrm{d}v \cdot v^3 e^{-2v/7} \ , \tag{4.308}$$

where we have once again used the change of variables $v = \gamma s/\log^2(\gamma s)$. We can evaluate the integral to give

$$\int_t^\infty \mathrm{d}s \tilde{w}_\gamma(s) \leq 46000e^2\gamma^{-1} \left[e^{-2v/7}(4v^3+42v^2+294v+1029)\right]_{v=\frac{\gamma t}{\log^2(\gamma t)}} \tag{4.309}$$

$$\leq 322000e^2\gamma^{-1} \left[e^{-2v/7}v^3\right]_{v=\frac{\gamma t}{\log^2(\gamma t)}} \tag{4.310}$$

$$= W_2\gamma^{-1} \left(\frac{\gamma t}{\log^2(\gamma t)}\right)^3 \exp\left(\frac{2\gamma t}{7\log^2(\gamma t)}\right) \ , \tag{4.311}$$

where the second inequality holds for $\gamma t \geq e^3$, and $W_2 = 322000e^2 \approx 2.38 \times 10^6$.

Finally, applying the previous bound, we see that

$$\int_{8500\gamma^{-1}}^\infty \mathrm{d}s \tilde{w}_\gamma(s) \leq 2^{-1}\gamma^{-1} \ . \tag{4.312}$$

So we can write

$$\int_{-\infty}^{\infty} \mathrm{d}t\tilde{w}_\gamma(t) = 2\int_0^{8500\gamma^{-1}} \mathrm{d}t\tilde{w}_\gamma(t) + 2\int_{8500\gamma^{-1}}^{\infty} \mathrm{d}t\tilde{w}_\gamma(t) \tag{4.313}$$

$$\leq 8501\gamma^{-1} = W_3\gamma^{-1} \ , \tag{4.314}$$

where we defined $W_3 = 8501$. ∎

Local truncations

In this section, we prove Lemma 4.2.15. This establishes that, for a local observable O_A supported on $A \subseteq \Gamma$, the operators $\Phi_H(O_A)$ and $\Psi_H(O_A)$ can both be well-approximated by operators which act only within a radius $r \geq 0$ of A, up to an error which decays quickly with r.

We first prove the following lemma, which shows that operators $\Lambda(O_A)$, of a form which generalises the quantum belief propagation and spectral flow operators, can be locally truncated. This will then be specialised to give the proof of Lemma 4.2.15.

Lemma 4.6.3. *Let H be a Hamiltonian on Γ with bounded interactions with respect to the $\|\cdot\|_{F_g}$ norm as in Lemma 4.2.14. Let O_A be a local observable supported on $A \subseteq \Gamma$, and let $w : \mathbb{R} \to \mathbb{R}$ be a non-negative L_1 function. We define the operator $\Lambda(O_A)$ (as a generalisation of the quantum belief propagation and spectral flow operators) by*

$$\Lambda(O_A) = \int_{-\infty}^{\infty} \mathrm{d}t w(t) e^{itH} O_A e^{-itH} \ , \tag{4.315}$$

for constants $c, \nu > 0$ depending on C_{F_g} and $\|F_g\|$. Then for any $r \geq 0$ there exists an operator $\Lambda^{[r]}(O_A)$ which acts as the identity outside of $B_r(A)$ such that

$$\|\Lambda(O_A) - \Lambda^{[r]}(O_A)\| \leq \|O_A\| \int_{-\infty}^{\infty} \mathrm{d}t w(t) \min\left\{2, c|A|\left(e^{\nu\|H\|_{F_g}|t|} - 1\right)e^{-g(r)}\right\} \ . \tag{4.316}$$

Moreover, assuming that $w(t)$ is even and normalised by $\int_{\mathbb{R}} \mathrm{d}t w(t) \leq 1$, the following bound holds:

$$\|\Lambda(O_A) - \Lambda^{[r]}(O_A)\| \leq \|O_A\| \inf_{\delta \in [0,\infty)} \left\{ c|A| e^{\nu\|H\|_{F_g}\delta - g(r)} + 4\int_{\delta}^{\infty} \mathrm{d}t w(t) \right\} \ . \tag{4.317}$$

To prove Lemma 4.6.3, we will define $\Lambda^{[r]}(O_A)$ simply by taking a partial trace on the sites outside $B_r(A)$, that is

$$\Lambda^{[r]}(O_A) := \frac{1}{\dim \mathcal{H}_{\Gamma \setminus B_r(A)}} \mathbb{1}_{\Gamma \setminus B_r(A)} \otimes \mathrm{tr}_{\Gamma \setminus B_r(A)}[\Lambda(O_A)] \ . \tag{4.318}$$

This is a slight abuse of notation; note that $\Lambda^{[r]}(O_A)$ depends on the choice of A, not just O_A.

Proof of Lemma 4.6.3. Letting $\mathcal{U}_{\Gamma\setminus B_r(A)}$ denote the Haar measure over unitaries on $\mathcal{H}_{\Gamma\setminus B_r(A)}$ (normalised as $\int_U \mathrm{d}U = 1$), we can write

$$\Lambda^{[r]}(O_A) = \int_{U\sim\mathcal{U}_{\Gamma\setminus B_r(A)}} \mathrm{d}U \left(U\otimes \mathbb{1}_{B_r(A)}\right)\Lambda(O_A)\left(U^\dagger\otimes \mathbb{1}_{B_r(A)}\right) \ . \tag{4.319}$$

We can then bound

$$\|\Lambda^{[r]}(O_A) - \Lambda(O_A)\| \leq \left\| \int_{U\sim\mathcal{U}_{\Gamma\setminus B_r(A)}} \mathrm{d}U \left[\left(U\otimes \mathbb{1}_{B_r(A)}\right)\Lambda(O_A)\left(U^\dagger\otimes \mathbb{1}_{B_r(A)}\right) - \Lambda(O_A)\right]\right\| \tag{4.320}$$

$$\leq \sup_U \left\|[U\otimes \mathbb{1}_{B_r(A)}, \Lambda(O_A)]\right\| \ . \tag{4.321}$$

By the definition of $\Lambda(O_A)$ we have

$$\left\|[U\otimes \mathbb{1}_{B_r(A)}, \Lambda(O_A)]\right\| \leq \int_{-\infty}^{\infty} \mathrm{d}t w(t) \left\|[U\otimes \mathbb{1}_{B_r(A)}, e^{itH}O_A e^{-itH}]\right\| \ . \tag{4.322}$$

Applying Lemma 4.2.14, we can bound the commutator by

$$\left\|[U\otimes \mathbb{1}_{B_r(A)}, e^{itH}O_A e^{-itH}]\right\| \leq c\|O_A\||A|(e^{\nu\|H\|_{F_g}|t|}-1)e^{-g(r)} \ . \tag{4.323}$$

Since U is unitary, we can also upper bound this by $2\|O_A\|$. Putting this together, we arrive at the bound

$$\|\Lambda(O_A) - \Lambda^{[r]}(O_A)\| \leq \|O_A\| \int_{-\infty}^{\infty} \mathrm{d}t w(t) \min\left\{2, c|A|\left(e^{\nu\|H\|_{F_g}|t|}-1\right)e^{-g(r)}\right\} \ . \tag{4.324}$$

Assuming that $w(t)$ is an even function, we can take the integral only over the positive real line, picking up a factor of two. Moreover, for any $\delta \geq 0$, we can split this integral into two parts: the $[0,\delta)$ interval, and the $[\delta,\infty)$ interval. This leads to the following bound:

$$\|\Lambda(O_A) - \Lambda^{[r]}(O_A)\| \leq 2\|O_A\| \left(\int_0^\delta \mathrm{d}t w(t) c|A|\left(e^{\nu\|H\|_{F_g}t}-1\right)e^{-g(r)} + \int_\delta^\infty \mathrm{d}t\cdot 2w(t)\right) \ . \tag{4.325}$$

Using the normalisation of $w(t)$, we can upper bound the first integral to obtain

$$\|\Lambda(O_A) - \Lambda^{[r]}(O_A)\| \leq \|O_A\| \left(c|A|e^{\nu\|H\|_{F_g}\delta - g(r)} + 4\int_\delta^\infty \mathrm{d}t w(t)\right) \ , \tag{4.326}$$

from which the result follows. ∎

We now prove Lemma 4.2.15, restated below.

Lemma 4.6.4 (Restatement of Lemma 4.2.15)**.** *Let H be a Hamiltonian with bounded $\|\cdot\|_{F_g}$-norm, and let O_A be an observable supported on $A \subseteq \Gamma$. Let $\Phi_H(O_A)$ be the quantum belief propagation operator defined by Eq.* (4.29)*, and $\Psi_H(O_A)$ be the spectral flow operator defined by Eq.* (4.35)*. Then for every $r \geq 0$ there exist operators $\Phi_H^{[r]}(O_A)$ and $\Psi_H^{[r]}(O_A)$ which act as the identity outside of $B_r(A)$, and constants $a_1, b_1 > 0$ (depending on $c, \beta, \nu, \|H\|_{F_g}$) and $a_2, b_2 > 0$ (depending on $c, \gamma, \nu, \|H\|_{F_g}$), such that*

$$\|\Phi_H(O_A) - \Phi_H^{[r]}(O_A)\| \leq a_1 |A| \|O_A\| e^{-b_1 g(r)} \ , \tag{4.327}$$

$$\|\Psi_H(O_A) - \Psi_H^{[r]}(O_A)\| \leq a_2 |A| \|O_A\| e^{-b_2 g(r)/\log^2 g(r)} \ , \tag{4.328}$$

where $\omega(x)$ is defined by

$$\omega(x) := \left(\frac{x}{\log^2(x)}\right)^3 \exp\left(-\frac{2x}{7\log^2 x}\right) \ . \tag{4.329}$$

Proof of Lemma 4.6.4. Application of Lemma 4.6.3 immediately yields the operator $\Phi_H^{[r]}(O_A)$, with the bound

$$\|\Phi_H(O_A) - \Phi_H^{[r]}(O_A)\| \leq \inf_{\delta \in [0,\infty)} \|O_A\| \left(c|A| e^{\nu \|H\|_{F_g} \delta - g(r)} + 4 \int_\delta^\infty \mathrm{d}t \kappa_\beta(t) \right) \tag{4.330}$$

where $\kappa_\beta(t)$ is the function defined in Eq. (4.30). By the bound given in Eq. (4.30) (see Ref. [193]), we have for $t \geq 0$ that

$$\kappa_\beta(t) \leq \frac{1}{e^{\frac{\pi t}{\beta}} - 1} = e^{-\frac{\pi t}{2\beta}} \cdot \frac{1}{2\sinh\left(\frac{\pi t}{2\beta}\right)} \ . \tag{4.331}$$

In particular, choosing $\delta \geq \delta_*$ where

$$\delta_* = \frac{2\beta}{\pi} \sinh^{-1}(1) \ , \tag{4.332}$$

we have

$$\int_\delta^\infty \mathrm{d}t \kappa_\beta(t) \leq \int_\delta^\infty \mathrm{d}t \frac{1}{2} e^{-\frac{\pi t}{2\beta}} = \frac{\beta}{\pi} e^{-\frac{\pi \delta}{2\beta}} \ . \tag{4.333}$$

Hence we can write

$$\|\Phi_H(O_A) - \Phi_H^{[r]}(O_A)\| \leq \inf_{\delta \in [\delta_*,\infty)} \|O_A\| \left(c|A| e^{\nu \|H\|_{F_g} \delta - g(r)} + \frac{4\beta}{\pi} e^{-\frac{\pi \delta}{2\beta}} \right) \tag{4.334}$$

$$\leq \|O_A\| \left(c|A| e^{\frac{2\beta}{\pi} \nu \|H\|_{F_g} \sinh^{-1}(1)} + \frac{4\beta}{\pi} \right) \exp\left[-\frac{\pi g(r)}{2\beta\nu\|H\|_{F_g} + \pi} \right] \ , \tag{4.335}$$

where in the last inequality we have set

$$\delta = \delta_* + \frac{g(r)}{\nu\|H\|_{F_g} + \frac{\pi}{2\beta}} \tag{4.336}$$

for an upper bound.

For the spectral flow, we may apply Lemma 4.2.10 and Lemma 4.6.3 to obtain $\Psi_H^{[r]}(O_A)$ which satisfies

$$\|\Psi_H(O_A) - \Psi_H^{[r]}(O_A)\| \leq \|O_A\| \inf_{\delta \in [e^3\gamma^{-1}, \infty)} \left\{ c|A|W_3\gamma^{-1} e^{\nu\|H\|_{F_g}\delta - g(r)} \right. \tag{4.337}$$

$$\left. + 4W_2\gamma^{-1} \left(\frac{\gamma\delta}{\log^2(\gamma\delta)} \right)^3 \exp\left(-\frac{2\gamma\delta}{7\log^2(\gamma\delta)} \right) \right\}, \tag{4.338}$$

where $W_2, W_3 > 0$ are the constants from Lemma 4.2.10. We define the shorthand

$$\omega(x) := \left(\frac{x}{\log^2 x} \right)^3 \exp\left(-\frac{2x}{7\log^2 x} \right) \geq e^{-2x/7}, \tag{4.339}$$

where the inequality holds for $x \geq e^3$. Choosing

$$\delta = e^3\gamma^{-1} + \frac{g(r)}{\nu\|H\|_{F_g} + \frac{2}{7}\gamma}, \tag{4.340}$$

we therefore obtain the upper bound

$$\|\Psi_H(O_A) - \Psi_H^{[r]}(O_A)\| \leq \|O_A\|\gamma^{-1} \left(c|A|W_3 e^{\nu\|H\|_{F_g} e^3\gamma^{-1}} + 4W_2 \right) \omega\left(\frac{g(r)}{\nu\gamma^{-1}\|H\|_{F_g} + \frac{2}{7}} \right), \tag{4.341}$$

as required. ∎

Generalised approximate local indistinguishability

We now apply the results of the previous section to establish the GALI property for parametrised Gibbs states and ground states, stated in the main text as Lemmas 4.2.16 and 4.2.17. These arguments are essentially the same as those in Ref. [193].

Lemma 4.6.5 (Restatement of Lemma 4.2.16)**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ be a family of Hamiltonians with continuous first derivative with respect to x on the interval $\mathcal{X}$, such that $\|H\|_{F_g}$ and $\|\partial_x H\|_{F_g}$ are both bounded. Let O_A be an observable supported on $A \subseteq \Gamma$, and assume that $\partial_s H(s)$ contains no terms with support in $B_{r_0}(A)$ for some $r_0 \geq 0$.*

Let $\rho_\beta(x)$ be the associated family of Gibbs states at temperature $\beta = \mathcal{O}(1)$, and assume that these satisfy a uniform exponential decay of correlations with parameters

$K, \xi > 0$. *Then, for all* x, *it holds that*

$$\left|\mathrm{Cov}_{\rho_\beta(x)}(\Phi_{H(x)}(O_A), \partial_x H(x))\right| \leq c_1 \|O_A\| |A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)} \right) , \tag{4.342}$$

for positive constants $b_1, c_1 > 0$. *Hence, defining* $f_\beta(x) := \mathrm{tr}[O_A \rho_\beta(x)]$, *we have for all* $x_0, x_1 \in \mathcal{X}$ *that*

$$|f_\beta(x_1) - f_\beta(x_0)| \leq \beta |x_1 - x_0| c_1 \|O_A\| |A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)} \right) . \tag{4.343}$$

Proof of Lemma 4.6.5. In the below, we will suppress some of the x-dependences for convenience. We write

$$\partial_x H = \sum_{r=r_0}^{\infty} \partial_x H^{[r]} , \tag{4.344}$$

where each $H^{[r]}$ contains the terms in H whose support intersects with $B_r(A)$, but not with $B_{r'}(A)$ for any $r' < r$. Using the triangle inequality, we can thus write

$$\left|\mathrm{Cov}_{\rho_\beta}(\Phi_H(O_A), \partial_x H)\right| \leq \sum_{r=r_0}^{\infty} \left|\mathrm{Cov}_{\rho_\beta}(\Phi_H(O_A), \partial_x H^{[r]})\right| . \tag{4.345}$$

For each $r \geq r_0$, we can decompose

$$\Phi_H(O_A) = \Phi_H^{[\lfloor r/2 \rfloor]}(O_A) + (\Phi_H(O_A) - \Phi_H^{[\lfloor r/2 \rfloor]}(O_A)) , \tag{4.346}$$

where by Lemma 4.2.15 the operator $\Phi_H^{[\lfloor r/2 \rfloor]}(O_A)$ is supported only within radius $\lfloor r/2 \rfloor$ of A, and

$$\|\Phi_H(O_A) - \Phi_H^{[\lfloor r/2 \rfloor]}(O_A)\| \leq a_1 |A| \|O_A\| e^{-b_1 g(r/2)} . \tag{4.347}$$

Hence, by the triangle inequality

$$\left|\mathrm{Cov}_{\rho_\beta}(\Phi_H(O_A), \partial_x H^{[r]})\right| \leq \left|\mathrm{Cov}_{\rho_\beta}(\Phi^{[\lfloor r/2 \rfloor]}(O_A), \partial_x H^{[r]})\right| + 2a_1 |A| \|O_A\| e^{-b_1 g(r/2)} \|\partial_x H^{[r]}\| \tag{4.348}$$

$$\leq 2\|O_A\| \|\partial_x H^{[r]}\| |B_r(A)|^2 K e^{-r/2\xi} + 2a_1 |A| \|O_A\| e^{-b_1 g(r/2)} \|\partial_x H^{[r]}\| \tag{4.349}$$

$$\leq 2\|O_A\| \|\partial_x H^{[r]}\| |A| \left(|A| K k_D^2 r^{2D} e^{-r/2\xi} + a_1 e^{-b_1 g(r/2)} \right) , \tag{4.350}$$

where in the second line we used the assumption of exponential correlation decay. Using Lemma 4.2.13, we can bound

$$\|\partial_x H^{[r]}\| \leq |B_r(A)| \|\partial_x H\|_{F_g} \|F_g\| \leq k_D r^D |A| \|\partial_x H\|_{F_g} \|F_g\| \ , \tag{4.351}$$

and hence we can write

$$\left|\mathrm{Cov}_{\rho_\beta}(\Phi_H(O_A), \partial_x H^{[r]})\right| \leq 2\|O_A\| k_D |A|^3 \|\partial_x H\|_{F_g} \|F_g\| (k_D^2 K + a_1) \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)}\right) \tag{4.352}$$

$$= c_1 \|O_A\| |A|^3 \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)}\right) \ , \tag{4.353}$$

where we have defined the constant

$$c_1 := 2k_D \|\partial_x H\|_{F_g} \|F_g\| (k_D^2 K + a_1) = \mathcal{O}(1) \ . \tag{4.354}$$

Applying Eq. (4.345), we can bound the overall covariance by

$$\left|\mathrm{Cov}_{\rho_\beta}(\Phi_H(O_A), \partial_x H)\right| \leq c_1 \|O_A\| |A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)}\right) \ . \tag{4.355}$$

Defining $f_\beta(x) := \mathrm{tr}[O_A \rho_\beta(x)]$, we can apply Eq. (4.32) to immediately deduce that

$$|f_\beta(x_1) - f_\beta(x_0)| \leq \int_{x_0}^{x_1} \mathrm{d}x \left|\frac{\mathrm{d}}{\mathrm{d}x} f_\beta(x)\right| \tag{4.356}$$

$$\leq \beta \int_{x_0}^{x_1} \mathrm{d}x \left|\mathrm{Cov}_{\rho_\beta(x)}(\Phi_{H(x)}(O_A), \partial_x H(x))\right| \tag{4.357}$$

$$\leq \beta |x_1 - x_0| c_1 \|O_A\| |A|^3 \sum_{r=r_0}^{\infty} \left(r^{3D} e^{-r/2\xi} + r^D e^{-b_1 g(r/2)}\right) \ , \tag{4.358}$$

as required. ∎

Lemma 4.6.6 (Restatement of Lemma 4.2.17)**.** *Let $\{H(x)\}_{x\in\mathcal{X}}$ and O_A be as in Lemma 4.6.5, and assume that $\{H(x)\}_{x\in\mathcal{X}}$ has a uniform gap $\gamma > 0$ above its ground state $|\psi_0(x)\rangle$. Then*

$$\left\|[\Psi_{H(x)}(O_A), \partial_x H(x)]\right\| \leq c_2 \|O_A\| |A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} \ , \tag{4.359}$$

for positive constants $b_2, c_2 > 0$. Hence, defining $f_{\mathsf{ground}}(x) := \langle\psi_0(x)|O_A|\psi_0(x)\rangle$, we have for all $x_0, x_1 \in \mathcal{X}$ that

$$|f_{\mathsf{ground}}(x_1) - f_{\mathsf{ground}}(x_0)| \leq |x_1 - x_0| c_2 \|O_A\| |A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} \ . \tag{4.360}$$

Proof of Lemma 4.6.6. We begin by defining $H^{[r]}$ as in Eq. (4.344), and use the triangle inequality to bound

$$\|[\Psi_H(O_A), \partial_x H]\| \leq \sum_{r=r_0}^{\infty} \left\| [\Psi_H(O_A), \partial_x H^{[r]}] \right\| \ . \tag{4.361}$$

For each $r \geq r_0$, we decompose

$$\Psi_H(O_A) = \Psi_H^{[r]}(O_A) + (\Psi_H(O_A) - \Psi_H^{[r]}(O_A)) \ , \tag{4.362}$$

where each $\Psi_H^{[r]}(O_A)$ is supported only within a radius r of A, and according to Lemma 4.2.15 we have

$$\|\Psi_H(O_A) - \Psi_H^{[r]}(O_A)\| \leq a_2 |A| \|O_A\| e^{-b_2 g(r)/\log^2 g(r)} \ . \tag{4.363}$$

Note that the operators $\Psi_H^{[r-1]}(O_A)$ and $\partial_s H^{[r]}$ commute as they have disjoint support, so we can bound

$$\begin{aligned}
\left\| [\Psi_H(O_A), \partial_x H^{[r]}] \right\| &\leq 2\|\Psi_H(O_A) - \Psi_H^{[r-1]}(O_A)\| \cdot \|\partial_x H^{[r]}\| && (4.364)\\
&\leq 2|A|\|O_A\| e^{-b_2 g(r-1)/\log^2 g(r-1)} \cdot k_D r^D |A| \|\partial_x H\|_{F_g} \|F_g\| && (4.365)\\
&= c_2 \|O_A\| |A|^2 r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} \ , && (4.366)
\end{aligned}$$

where in the second line we have used Eqs. (4.351) and (4.363), and where

$$c_2 := 2k_D \|\partial_x H\|_{F_g} \|F_g\| \ . \tag{4.367}$$

Using Eq. (4.361), we arrive at the required bound:

$$\|[\Psi_H(O_A), \partial_x H]\| \leq c_2 \|O_A\| |A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} \ . \tag{4.368}$$

Defining $f_{\mathsf{ground}}(x) := \langle \psi_0(x) | O_A | \psi_0(x) \rangle$, we can apply Eq. (4.37) to deduce that

$$\begin{aligned}
|f_{\mathsf{ground}}(x_1) - f_{\mathsf{ground}}(x_0)| &\leq \int_{x_0}^{x_1} \mathrm{d}x \left| \frac{\mathrm{d}}{\mathrm{d}x} f_{\mathsf{ground}}(x) \right| && (4.369)\\
&\leq \int_{x_0}^{x_1} \|[\Psi_H(O_A), \partial_x H]\| && (4.370)\\
&\leq |x_1 - x_0| c_2 \|O_A\| |A|^2 \sum_{r=r_0}^{\infty} r^D e^{-b_2 g(r-1)/\log^2 g(r-1)} \ , && (4.371)
\end{aligned}$$

as required. ∎

4.6.3 Bounding the local Schrieffer-Wolff expansion

Preliminary results

In this section, we prove a couple of lemmas controlling the growth of the $\|\cdot\|_F$-norm under commutator, which will ultimately be important to bound the terms of the perturbative expansion $V^{(q)}$ obtained via the local Schrieffer-Wolff transformation.

Lemma 4.6.7. *Let F be a normalised F-function and let $X = \sum_{A\subseteq\Gamma} X_A$ and $Y = \sum_{A\subseteq\Gamma} Y_A$ be Hamiltonians on Γ. Then*

$$\|[X,Y]\|_F \leq 4\|X\|_F\|Y\|_F \ . \tag{4.372}$$

Proof of Lemma 4.6.7. Writing $[X,Y]$ as a sum of local terms, $[X,Y] = \sum_{A\subseteq\Gamma}[X,Y]_A$, by definition we have

$$\|[X,Y]\|_F \leq \sup_{i,j\in\Gamma} \frac{1}{F(\mathrm{dist}(i,j))} \sum_{\substack{A\subseteq\Gamma \\ i,j\in A}} \|[X,Y]_A\| \ . \tag{4.373}$$

But for fixed i,j, note that nonzero terms $[X,Y]_A$ can only arise from as commutators between terms X_B and Y_C with overlapping support (containing at least one $k \in B\cap C$), i.e.

$$\sum_{\substack{A\subseteq\Gamma \\ i,j\in A}} \|[X,Y]_A\| \leq \sum_{k\in\Gamma} \left(\sum_{\substack{B\ni i,k \\ C\ni j,k}} (\|[X_B,Y_C]\| + \|[X_C,Y_B]\|) \right) \tag{4.374}$$

$$\leq 2\sum_{k\in\Gamma} \left(\left(\sum_{B\ni i,k} \|X_B\| \right) \left(\sum_{C\ni j,k} \|Y_C\| \right) + \left(\sum_{B\ni i,k} \|Y_B\| \right) \left(\sum_{C\ni j,k} \|X_C\| \right) \right) \ . \tag{4.375}$$

But we have $\sum_{A\ni i,k} \|X_A\| \leq F(\mathrm{dist}(i,k))\|X\|_F$, and similarly for the other terms, so the above can be bounded by

$$\sum_{\substack{A\subseteq\Gamma \\ i,j\in A}} \|[X,Y]_A\| \leq 4\|X\|_F\|Y\|_F \sum_{k\in\Gamma} F(\mathrm{dist}(i,k))F(\mathrm{dist}(j,k)) \ . \tag{4.376}$$

Thus

$$\|[X,Y]\|_F \leq 4\|X\|_F\|Y\|_F \sup_{i,j\in\Gamma} \sum_{k\in\Gamma} \frac{F(\mathrm{dist}(i,k))F(\mathrm{dist}(j,k))}{F(\mathrm{dist}(i,j))} \tag{4.377}$$

$$\leq 4\|X\|_F\|Y\|_F \ , \tag{4.378}$$

as required, where we have used the normalisation condition (the constant C_F defined in Eq. (4.44) is equal to 1). ∎

Lemma 4.6.8. *Let W be a Hamiltonian, and $q \geq 1$. Then*

$$\|[T^{(q)}, W]\|_F \leq 4\Delta^{-1}\|V^{(q)}\|_F\|W\|_F \ . \tag{4.379}$$

Proof of Lemma 4.6.8. Using Lemma 4.6.7, we have

$$\|[T^{(q)}, W]\|_F \leq 4\|T^{(q)}\|_F\|W\|_F \ . \tag{4.380}$$

Moreover, using the definition Eq. (4.175) with Lemma 4.4.1(I), we have

$$\|T^{(q)}\|_F \leq \Delta^{-1}\|V^{(q)}\|_F \ , \tag{4.381}$$

and the result follows. ∎

Bounding series terms

Below we restate and prove Lemma 4.4.2, which places bounds on the $V^{(q)}$ constructed via the local Schrieffer-Wolff transformation. The proof is very similar to Ref. [172], Lemma 4.2, but differs in the respects that it accounts for general polynomial perturbations in x, and that Lemmas 4.6.7 and 4.6.8 provide tighter bounds for geometrically local Hamiltonians which leads to a provably convergent series.

Lemma 4.6.9 (Restatement of Lemma 4.4.2)**.** *For $q \geq 1$, $V^{(q)}$ is bounded as*

$$\|V^{(q)}\|_F \leq \frac{\Delta\theta^q}{16} \ , \tag{4.382}$$

where $\theta > 0$ is a constant depending on the ratio J/Δ. In particular, if $x \leq 1/(2\theta)$, then the Hamiltonian $V(x) = \sum_{q\geq 1} x^q V^{(q)}$ is bounded as

$$\|V_{\mathsf{eff}}(x)\|_F \leq \|V(x)\|_F \leq \frac{\Delta\theta x}{8} \ . \tag{4.383}$$

Proof of Lemma 4.6.9. By definition, we have $\|V^{(1)}\|_F = \|H^{(1)}\|_F \leq J$. Using Eq. (4.161) and the triangle inequality, we have the following recursive bound for $q \geq 2$:

$$\begin{aligned}\|V^{(q)}\|_F \leq \|H^{(q)}\|_F &+ \sum_{r=2}^{q} \frac{1}{r!} \sum_{\substack{q-1\geq q_1,\ldots,q_r\geq 1 \\ q_1+\cdots+q_r=q}} \|[T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(0)}]\cdots]]\| \\ &+ \sum_{\alpha=1}^{q}\sum_{r=1}^{q-\alpha} \frac{1}{r!} \sum_{\substack{q-\alpha\geq q_1,\ldots,q_r\geq 1 \\ q_1+\cdots+q_r=q-\alpha}} \|[T^{(q_1)}, [T^{(q_2)}, \ldots [T^{(q_r)}, H^{(\alpha)}]\cdots]]\| \ . \end{aligned} \tag{4.384}$$

By assumption, $\|H^{(q)}\|_F \leq J$. Moreover, using Lemma 4.6.8 r times we can bound

$$\|[T^{(q_1)}, \dots [T^{(q_r)}, H^{(\alpha)}] \dots]\| \leq (4\Delta^{-1})^r \|V^{(q_1)}\|_F \dots \|V^{(q_r)}\|_F J \ , \quad \alpha \geq 0 \ . \tag{4.385}$$

Hence we have

$$\begin{aligned} \|V^{(q)}\|_F \leq J &+ \sum_{r=2}^{q} \frac{1}{r!} \sum_{\substack{q-1 \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q}} (4\Delta^{-1})^r \|V^{(q_1)}\|_F \dots \|V^{(q_r)}\|_F J \\ &+ \sum_{\alpha=1}^{q} \sum_{r=1}^{q-\alpha} \frac{1}{r!} \sum_{\substack{q-\alpha \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q-\alpha}} (4\Delta^{-1})^r \|V^{(q_1)}\|_F \dots \|V^{(q_r)}\|_F J \end{aligned} \tag{4.386}$$

$$\Rightarrow \frac{\|V^{(q)}\|_F}{J} \leq 1 + \sum_{r=2}^{q} c^r \sum_{\substack{q-1 \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q}} \prod_{i=1}^{r} \left(\frac{\|V^{(q_i)}\|_F}{J} \right) + \sum_{\alpha=1}^{q} \sum_{r=1}^{q-\alpha} c^r \sum_{\substack{q-\alpha \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q-\alpha}} \prod_{i=1}^{r} \left(\frac{\|V^{(q_i)}\|_F}{J} \right) \ . \tag{4.387}$$

where we have defined the constant $c = 4J/\Delta$ and generously lower bounded $r! \geq 1$.

We can use this expression to recursively define a sequence $\{\mu_q\}_{q=1}^{\infty}$ such that $\|V^{(q)}\|_F / J \leq \mu_q$ for all $q \geq 1$. We set $\mu_1 = 1$ and

$$\mu_q = 1 + \sum_{r=2}^{q} c^r \sum_{\substack{q-1 \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q}} \mu_{q_1} \dots \mu_{q_r} + \sum_{\alpha=1}^{q-1} \sum_{r=1}^{q-\alpha} c^r \sum_{\substack{q-\alpha \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q-\alpha}} \mu_{q_1} \dots \mu_{q_r} \ , \quad q \geq 2 \ , \tag{4.388}$$

which, after some algebra, can be rearranged to

$$(1+c)\mu_q = 1 + \sum_{\alpha=0}^{q-1} \sum_{r=1}^{q-\alpha} c^r \sum_{\substack{q-\alpha \geq q_1, \dots, q_r \geq 1 \\ q_1 + \dots + q_r = q-\alpha}} \mu_{q_1} \dots \mu_{q_r} \ , \quad q \geq 2 \ . \tag{4.389}$$

To make further progress, we define the function $\mu(z)$ as a formal power series

$$\mu(z) := \sum_{q=1}^{\infty} \mu_q z^q \ . \tag{4.390}$$

Notice that $\sum_{q\geq 1} z^q = z/(1-z)$ and

$$\sum_{q=2}^{\infty} z^q \sum_{\alpha=0}^{q-1}\sum_{r=1}^{q-\alpha} c^r \sum_{\substack{q-\alpha\geq q_1,\ldots,q_r\geq 1\\ q_1+\cdots+q_r=q-\alpha}} \mu_{q_1}\ldots\mu_{q_r} = \sum_{q=0}^{\infty} z^q \sum_{\alpha=0}^{q-1}\sum_{r=1}^{q-\alpha} c^r \sum_{\substack{q-\alpha\geq q_1,\ldots,q_r\geq 1\\ q_1+\cdots+q_r=q-\alpha}} \mu_{q_1}\ldots\mu_{q_r} - cz \tag{4.391}$$

$$= \sum_{\alpha=0}^{\infty} z^{\alpha} \sum_{r=1}^{\infty} c^r \sum_{q_1,\ldots,q_r=1}^{\infty} z^{q_1+\cdots+q_r}\mu_{q_1}\ldots\mu_{q_r} - cz \tag{4.392}$$

$$= \frac{1}{1-z}\cdot\frac{c\mu(z)}{1-c\mu(z)} - cz\ . \tag{4.393}$$

Thus, the condition that $\mu_1 = 1$ and the recursion relation Eq. (4.389) are equivalent to

$$(1+c)\mu(z) = \frac{z}{1-z} + \frac{1}{1-z}\cdot\frac{c\mu(z)}{1-c\mu(z)}\ , \tag{4.394}$$

which can be rearranged to

$$c(1+c)\mu(z)^2 - \mu(z) + \frac{z}{1-z} = 0\ . \tag{4.395}$$

Solving this quadratic, and choosing the branch with $\mu(0) = 0$, we obtain

$$\mu(z) = \frac{1}{2c(1+c)}\left(1 - \sqrt{1 - 4c(1+c)\frac{z}{1-z}}\right)\ . \tag{4.396}$$

This function is analytic in the disc

$$|z| < z_0 = \frac{1}{4c(1+c)+1}\ , \tag{4.397}$$

and moreover, defining $z_1 = 1/(8c(1+c)+1)$, we have

$$|\mu(z)| \leq \frac{1}{2c(1+c)}\left(1 - \frac{1}{\sqrt{2}}\right) \leq \frac{1}{4c(1+c)}\ , \quad \text{for } |z|\leq z_1\ . \tag{4.398}$$

Hence, using Cauchy's integral formula, we can bound the coefficients μ_q as

$$\mu_q = \frac{1}{2\pi i}\oint_{|z|=z_1} \frac{\mu(z)}{z^{q+1}}\,\mathrm{d}z \tag{4.399}$$

$$\leq z_1^{-q} \sup_{|z|=z_1} |\mu(z)| \tag{4.400}$$

$$\leq (8c(1+c)+1)^q \cdot \frac{1}{4c(c+1)} \tag{4.401}$$

$$\leq \frac{(8(1+c))^{2q}}{4c}\ . \tag{4.402}$$

Hence

$$\|V^{(q)}\|_F \leq J\mu_q \leq J \cdot \frac{(8(1+c))^{2q}}{4c} \tag{4.403}$$

$$= \frac{\Delta}{16}\left(8\left(1+\frac{4J}{\Delta}\right)\right)^{2q} \tag{4.404}$$

$$= \frac{\Delta\theta^q}{16} , \tag{4.405}$$

where

$$\theta = 64(1+4J/\Delta)^2 . \tag{4.406}$$

It follows by the triangle inequality that

$$\|V(x)\|_F \leq \sum_{q\geq 1} x^q \|V^{(q)}\|_F \tag{4.407}$$

$$\leq \frac{\Delta}{16}\sum_{q\geq 1}(\theta x)^q \tag{4.408}$$

$$= \frac{\Delta\theta x}{16}\sum_{q\geq 0}(\theta x)^q \leq \frac{\Delta\theta x}{8} , \tag{4.409}$$

where the last inequality holds if $x \leq 1/(2\theta)$. ∎

4.6.4 Properties of Hamiltonian gadgets

Equivalent characterisation

In this section, we show that Definition 4.4.7 is equivalent in the high-energy regime to an alternative general formulation for gadgets introduced in a previous work (Ref. [1]), restated below. The formulation of Ref. [1] is itself based on notions of simulation from prior works Refs. [87, 56], so this result can be viewed as a consistency check with these also.

Definition 4.6.10 ((Δ,η,ϵ)-gadget — [1], Definition 18)**.** Let H' be a Hamiltonian acting on $\mathcal{H}' = \bigotimes_{i\in\Gamma_{\mathsf{eff}}\cup\Gamma_{\mathsf{anc}}}\mathcal{H}_i$, and let H be a Hamiltonian on $\mathcal{H}_{\mathsf{eff}} = \bigotimes_{i\in\Gamma_{\mathsf{eff}}}\mathcal{H}_i$. For constants $\Delta,\eta,\epsilon \geq 0$, we say that H' is a (Δ,η,ϵ)-gadget for H if there exists a rank-1 projector $P \in \mathrm{Proj}(\mathcal{H}_{\mathsf{anc}})$ and a unitary $U \in \mathrm{U}(\mathcal{H}')$ such that the projector onto subspace of states below energy Δ (with respect to H'), $P' \in \mathrm{Proj}(\mathcal{H}')$, can be written as $P' = U(\mathbb{1}\otimes P)U^\dagger$, and

$$\|U - \mathbb{1}\| \leq \eta , \quad \|P'H'P' - U(H\otimes P)U^\dagger\| \leq \epsilon . \tag{4.410}$$

This definition defines gadgets entirely in terms of their behaviour in a low-energy subspace, and in particular only has meaning when H' applies energy penalties to the ancillary sites of the order of $\|H\|$. Definition 4.4.7 only requires H' to be locally gapped, however in the limit $x \to 0$ these definitions coincide as we formalise in the following lemma:

Lemma 4.6.11 (Equivalent characterisation of gadgets)**.** *Let $\{H'(x)\}_{x\in(0,\infty)}$ be a family of Hamiltonians on $\mathcal{H}' = \otimes_{i\in\Gamma_{\text{anc}}\cup\Gamma_{\text{eff}}}\mathcal{H}_i$ which is a degree-d polynomial in x^{-1} written as $H'(x) = x^{-d}\sum_{\alpha=0}^{d} x^\alpha H^{(\alpha)}$. As in Definition 4.4.7, we assume that $H^{(0)} = \Delta'\sum_{i\in\Gamma_{\text{anc}}}(\mathbb{1} - |0_i\rangle\langle 0_i|)$, where $\Delta' > 0$ and $|0_i\rangle$ is a local state on site i. Assume $\|H^{(\alpha)}\|_F \leq J$ for $1 \leq \alpha \leq d$, where F is a normalised F-function on Γ. The following are equivalent:*

(I) For sufficiently small x, $H'(x)$ is a $(\Delta(x), \eta(x), \epsilon(x))$-gadget for H, for some functions $\Delta, \eta, \epsilon : (0,\infty) \to [0,\infty)$ such that

$$\lim_{x\to 0}\Delta(x) = \infty \ , \quad \lim_{x\to 0}\eta(x) = 0 \ , \quad \lim_{x\to 0}\epsilon(x) = 0 \ . \tag{4.411}$$

(II) $H'(x)$ is a gadget of degree d for H.

This result essentially follows by noticing that, for sufficiently small x, the projection obtained through Schrieffer-Wolff perturbation theory will project onto the low-energy subspace of $H'(x)$. In this regime, Definitions 4.4.7 and 4.6.10 are in correspondence.

Proof of Lemma 4.6.11. The local Schrieffer-Wolff transformation produces series $\{V^{(q)}\}_{q\geq 1}$ and $\{T^{(q)}\}_{q\geq 1}$ such that

$$e^{T(x)}H'(x)e^{-T(x)} = x^{-d}H^{(0)} + x^{-d}\sum_{q\geq 1} x^q\left(V^{(q)} + [T^{(q)}, H^{(0)}]\right) \ , \tag{4.412}$$

where the sum is convergent for sufficiently small x (as given explicitly by Lemma 4.4.2), and all summands are block-diagonal with respect to the projector $P_0 = \otimes_{i\in\Gamma_{\text{anc}}}|0_i\rangle\langle 0_i|$. In particular, all states in the image of P_0 have energy at most $\mathcal{O}(x^{1-d})$, whilst all states in the image of $(1 - P_0)$ have energy at least $\Omega(x^{-d})$ due to the penalty imposed by the $H^{(0)}$ term. It follows that, for sufficiently small x, the projector $P' = e^{-T(x)}P_0e^{T(x)}$ projects onto the low-energy space of $H'(x)$ with cutoff $\Delta = \Theta(\Delta' x^{-d})$.

Suppose that $H'(x)$ is a gadget of degree d for H. This implies that the part of $e^{T(x)}H'(x)e^{-T(x)}$ lying in the P_0 space is a polynomial in x with constant term given by $H \otimes P_0$, and in particular we can write

$$\left\| P'H'(x)P' - e^{-T(x)}(H\otimes P_0)e^{T(x)}\right\| \leq \mathcal{O}(x) \ . \tag{4.413}$$

Hence, using that $\|T(x)\| = \mathcal{O}(x)$ and identifying $U = e^{-T(x)}$, we immediately recover Definition 4.6.10 as required.

Now suppose that $H'(x)$ is a $(\Delta(x), \eta(x), \epsilon(x))$-gadget for H, where $\Delta \to \infty$, $\eta, \epsilon \to 0$ as $x \to 0$. This guarantees that, for small x,

$$\|P'H'(x)P'\| \leq \|H\| + \mathcal{O}(x) \ , \tag{4.414}$$

which in particular implies that all negative powers of x must disappear in the P_0 space of $e^{T(x)}H'(x)e^{-T(x)}$. This, along with the observation that $P'H'(x)P' \to H \otimes P_0$ as $x \to 0$ (using that $\epsilon, \eta \to 0$), ensures that the constant term in the expansion is H, recovering Definition 4.4.7. ∎

Combination of gadgets

In this section, we prove some results concerning the combination of gadgets in parallel. Firstly, in the below lemma, we observe that a gadget can trivially be combined with a background Hamiltonian independent of the parameter x.

Lemma 4.6.12. *Suppose that $H'(x)$ is a gadget of degree d for H. Then, for any local Hamiltonian $\tilde{H}$ acting on $\mathcal{H}_{\mathsf{eff}}$, $H'(x) + \tilde{H}$ is a gadget of degree d for $H + \tilde{H}$.*

Proof of Lemma 4.6.12. This can be seen immediately from the construction of the perturbative series $\{V^{(q)}\}_{q\geq 1}$ as in Eq. (4.161). Notice that the constant term $\tilde{H}$ may be absorbed into $H^{(d)}$, which only appears at dth order in the perturbative expansion. ∎

It turns out that gadgets of degree $d \leq 3$ can immediately be used in parallel as guaranteed by the below theorem. Intuitively, this is because the interactions are not strong enough for cross-gadget contributions in the effective Hamiltonian, since this requires terms which excite and de-excite at least two different gadgets — such terms only appear at fourth order and above. Parallel combination results for gadgets have been proved in many specialised settings to prove results in Hamiltonian complexity theory, see e.g. Refs. [135, 56]. The below theorem can be viewed as a somewhat cleaner version of the previous general result of Ref. [1], Proposition 24.

Theorem 4.6.13 (Gadget combination)**.** *Let $\{h'_j(x)\}_{j\in\Gamma_{\mathsf{anc}}}$ be such that, for every $j \in \Gamma_{\mathsf{anc}}$, $h'_j(x)$ is a family of Hamiltonians acting on $\mathcal{H}_{\mathsf{eff}} \otimes \mathcal{H}_j$ which is a gadget of degree $d \leq 3$ for h_j on $\mathcal{H}_{\mathsf{eff}}$. Then the Hamiltonian*

$$H'(x) := \sum_{j\in\Gamma_{\mathsf{anc}}} h'_j(x) \ , \tag{4.415}$$

which acts on $\mathcal{H}' = \otimes_{x\in\Gamma_{\mathsf{eff}}\cup\Gamma_{\mathsf{anc}}} \mathcal{H}_x$, is a gadget of degree d for the combined Hamiltonian

$$H(x) := \sum_{j\in\Gamma_{\mathsf{anc}}} h_j \ , \tag{4.416}$$

on $\mathcal{H}_{\mathsf{eff}}$.

Since a $H'(x)$ constructed in this way is typically extensive, it is more natural to refer to this as a simulator Hamiltonian rather than a gadget as mentioned above. The proof of Theorem 4.6.13 follows by explicitly computing the first three terms $V^{(1)}, V^{(2)}, V^{(3)}$ of the local Schrieffer-Wolff expansion. We will require the following identity:

Lemma 4.6.14. *Let $A, B, C \subseteq \Gamma$ be disjoint sets of sites. Assume that X_{AB} and Y_{BC} are Hermitian operators supported only on the sites $A\cup B$ and $B\cup C$ respectively. Then*

$$P_{ABC}\left[\mathcal{L}_{AB}(X_{AB}), Y_{BC}\right] P_{ABC} = P_{ABC}\left[\mathcal{L}_{BC}(Y_{BC}), X_{AB}\right] P_{ABC} \ , \tag{4.417}$$

where the projector P_{ABC} and the superoperators $\mathcal{L}_{AB}, \mathcal{L}_{BC}$ are defined as in Eqs. (4.162) *and* (4.165) *respectively.*

Proof of Lemma 4.6.14. Using (4.165), we can write the left-hand side of the above expression as

$$\int_0^\infty \mathrm{d}t P_{ABC}\big[e^{-tH^{(0)}|_{AB}/\Delta}Q_{AB}X_{AB}P_{AB} - P_{AB}X_{AB}Q_{AB}e^{-tH^{(0)}|_{AB}/\Delta}, Y_{BC}\big]P_{ABC} \tag{4.418}$$

$$= -\int_0^\infty \mathrm{d}t\Big(P_{ABC}X_{AB}Q_{AB}e^{-tH^{(0)}|_{AB}/\Delta}Y_{BC}P_{ABC} + P_{ABC}Y_{BC}e^{-tH^{(0)}|_{AB}/\Delta}Q_{AB}X_{AB}P_{ABC}\Big) \tag{4.419}$$

$$= -\int_0^\infty \mathrm{d}t\Big(P_{ABC}X_{AB}Q_{B}e^{-tH^{(0)}|_{B}/\Delta}Y_{BC}P_{ABC} + P_{ABC}Y_{BC}e^{-tH^{(0)}|_{B}/\Delta}Q_{B}X_{AB}P_{ABC}\Big)\ . \tag{4.420}$$

In the last line, we have used that $H^{(0)}$ is 1-local to decompose $e^{-tH^{(0)}|_{AB}/\Delta} = e^{-tH^{(0)}|_{A}/\Delta}e^{-tH^{(0)}|_{B}/\Delta}$. In both terms, the part acting on A is projected into the P_A space on which it acts as the identity. Additionally, we have used the fact that $P_AQ_{AB} = P_AQ_B$. An identical calculation on the right-hand side of (4.417) arrives at the same expression. ∎

Proof of Theorem 4.6.13. Writing each gadget $h'_j(x)$ as a sum of terms $h'_j(x) = x^{-d}\sum_{\alpha=0}^d x^\alpha h_j^{(\alpha)}$, we can write $H'(x)$ as

$$H'(x) = x^{-d}\sum_{\alpha=0}^{d} x^\alpha H^{(\alpha)}\ , \quad \text{where} \quad H^{(\alpha)} := \sum_{j\in[N]} h_j^{(\alpha)}\ . \tag{4.421}$$

We will explicitly compute the first three terms $V^{(q)}$ of the local Schrieffer-Wolff transformation for this Hamiltonian, and show that their restriction to the $(\mathbb{1}\otimes P_0)$ space give to the correct target Hamiltonian.

For a local Hamiltonian $X = \sum_{A\subseteq\Gamma} X_A$, we will adopt the shorthand

$$\mathcal{L}(X) := \sum_{A\subseteq\Gamma}\mathcal{L}_A(X_A) \qquad \mathcal{O}_A(X) := \sum_{A\subseteq\Gamma}\mathcal{O}_A(X_A)\ . \tag{4.422}$$

Each of the Hamiltonians $V^{(1)}$, $V^{(2)}$, and $V^{(3)}$ will be a polynomial in the $h_j^{(\alpha)}$. Note that — for the restriction of $V^{(q)}$ to the $(\mathbb{1}\otimes P_0)$ subspace — we can ignore all of these polynomial terms which only contain one distinct j-index, because these terms will cancel due to our assumption that each of the $h'_j(x)$ are gadgets.

For the first-order term, we have

$$V^{(1)} = H^{(1)} = \sum_j h_j^{(1)}\ . \tag{4.423}$$

By assumption, $(\mathbb{1}\otimes P_0)h_j^{(1)}(\mathbb{1}\otimes P_0) = 0$ for all j, so this term will vanish. From (4.175) we have

$$T^{(1)} = \mathcal{L}(V^{(1)}) = \mathcal{L}(H^{(1)})\ . \tag{4.424}$$

For the second-order term, from (4.161) we have

$$V^{(2)} = H^{(2)} + \big[T^{(1)}, V^{(1)} - \frac{1}{2}\,\mathcal{O}(V^{(1)})\big] \tag{4.425}$$

$$= H^{(2)} + \big[\mathcal{L}(H^{(1)}), H^{(1)} - \frac{1}{2}\,\mathcal{O}(H^{(1)})\big] \tag{4.426}$$

$$= \sum_j \big(h_j^{(2)} + \big[\mathcal{L}(h_j^{(2)}), h_j^{(1)} - \frac{1}{2}\,\mathcal{O}(h_j^{(1)})\big] + \sum_{i\neq j}\big[\mathcal{L}(h_i^{(2)}), h_j^{(1)} - \frac{1}{2}\,\mathcal{O}(h_j^{(1)})\big]\ . \tag{4.427}$$

As mentioned, the first term can be ignored; by the assumption that each $h_j'(x)$ is a gadget, its contribution to the P_0 space will be zero. For the second term, notice that $\mathcal{L}(h_i^{(2)})$ is off-diagonal with respect to $P_{0,i} := |0_i\rangle\langle 0_i|$ by construction. However, by the fact that $[h_i^{(0)}, h_j^{(1)}] = 0$ (since $h_i^{(0)}$ is supported only on $\mathcal{H}_i$, on which $h_j^{(1)}$ does not act), we are guaranteed that the second term in the commutator is diagonal with respect to $P_{0,i}$. Hence the entire commutator is off-diagonal with respect to $P_{0,i}$, hence its contribution to the P_0 space will also be zero.

For notational convenience, we write

$$G := H^{(1)} - \frac{1}{2}\,\mathcal{O}(H^{(1)}) = \sum_j g_j\ , \quad g_j := h_j^{(1)} - \frac{1}{2}\,\mathcal{O}(h_j^{(1)})\ . \tag{4.428}$$

Then from (4.175) we have

$$V^{(2)} = H^{(2)} + [\mathcal{L}(H^{(1)}), G], \quad T^{(2)} = \mathcal{L}(H^{(2)}) + \mathcal{L}\big([\mathcal{L}(H^{(1)}), G]\big)\ . \tag{4.429}$$

For the third-order term, from (4.161) we then have

$$\begin{aligned} V^{(3)} &= H^{(3)} + \frac{1}{2}[T^{(1)}, -\mathcal{O}(V^{(2)})] + \frac{1}{2}[T^{(2)}, -\mathcal{O}(V^{(1)})] + \frac{1}{6}\big[T^{(1)}, [T^{(1)}, -\mathcal{O}(V^{(1)})]\big] \\ &\quad + [T^{(2)}, H^{(1)}] + \frac{1}{2}\big[T^{(1)}, [T^{(1)}, H^{(1)}]\big] + [T^{(1)}, H^{(2)}] \\ &= H^{(3)} + [\mathcal{L}(H^{(1)}), H^{(2)} - \frac{1}{2}\,\mathcal{O}(H^{(2)})] + [\mathcal{L}(H^{(2)}), G] && \text{(4.430a)} \\ &\quad + \big[\mathcal{L}([\mathcal{L}(H^{(1)}), G]), G\big] - \frac{1}{2}\big[\mathcal{L}(H^{(1)}), \mathcal{O}([\mathcal{L}(H^{(1)}), G])\big] && \text{(4.430b)} \\ &\quad + \frac{1}{2}\big[\mathcal{L}(H^{(1)}), [\mathcal{L}(H^{(1)}), H^{(1)}]\big] && \text{(4.430c)} \\ &\quad - \frac{1}{6}\big[\mathcal{L}(H^{(1)}), [\mathcal{L}(H^{(1)}), \mathcal{O}(H^{(1)})]\big] && \text{(4.430d)} \end{aligned}$$

We will now deal with the lines (4.430a)-(4.430d) separately.

- (4.430a):

  As mentioned above, we only need to worry about the terms consisting of contributions from two distinct gadgets; that is:

$$\sum_{i\neq j}\big([\mathcal{L}(h_i^{(1)}), h_j^{(2)} - \frac{1}{2}\,\mathcal{O}(h_j^{(2)})] + [\mathcal{L}(h_i^{(2)}), g_j]\big)\ . \tag{4.431}$$

For each term in this sum, notice that $\mathcal{L}(h_i^{(1)})$ and $\mathcal{L}(h_i^{(2)})$ are block-off-diagonal with respect to $P_{0,i}$ by definition, whilst $h_j^{(2)} - \frac{1}{2}\mathcal{O}(h_j^{(2)})$ and g_j commute with $h_i^{(0)}$, and are hence block-diagonal with respect to $P_{0,i}$. So the entire term is block-off-diagonal and has no contribution to the P_0 space.

- (4.430b):

  Notice that $\mathcal{L}$ only depends on the local off-diagonal parts of all its input terms (in other words, $\mathcal{L} = \mathcal{L} \circ \mathcal{O}$), and in particular $\mathcal{L}(H^{(1)}) = \frac{1}{2}\mathcal{L}(G)$. Hence we can write the terms in (4.430b) as

  $$2\big[\mathcal{L}(\mathcal{O}([\mathcal{L}(G), G])), G\big] - 2\big[\mathcal{L}(G), \mathcal{O}([\mathcal{L}(G), G])\big] \ . \tag{4.432}$$

  However, separating G into its local terms and applying Lemma 4.6.14, we see that these terms cancel in the P_0 space, giving zero contribution overall.

- (4.430c):

  Separating into individual gadgets, we are concerned with terms of the form

  $$\big[\mathcal{L}(h_i^{(1)}), [\mathcal{L}(h_j^{(1)}), h_k^{(1)}]\big] \ , \tag{4.433}$$

  where i, j, k are not all equal. Note that $\mathcal{L}(h_i^{(1)})$ and $\mathcal{L}(h_j^{(1)})$ are block off-diagonal with respect to $P_{0,i}$ and $P_{0,j}$ respectively. Moreover, due to the condition that the first-order term vanishes in the P_0 space, we know that $(\mathbb{1} \otimes P_{0,k}) h_k^{(1)} (\mathbb{1} \otimes P_{0,k}) = 0$. This means that, unless $i = k$ or $j = k$, the overall term will have no contribution to the $P_{0,k}$ space. If $i = k$, then $i \neq j$ by assumption, and the overall term is off-diagonal in the $P_{0,j}$ space. Likewise if $j = k$, then the overall term is off-diagonal in the $P_{0,i}$ space. Hence in all cases except $i = j = k$, this term provides no contribution to the P_0 space.

- (4.430d):

  For this term the argument is similar to the previous; this time we are interested in contributions of the form

  $$\big[\mathcal{L}(h_i^{(1)}), [\mathcal{L}(h_j^{(1)}), \mathcal{L}(h_k^{(1)})]\big] \ , \tag{4.434}$$

  where i, j, k are not all equal. Here, $\mathcal{L}(h_i^{(1)})$, $\mathcal{L}(h_j^{(1)})$, and $\mathcal{O}(h_k^{(1)})$ are off-diagonal with respect to $P_{0,i}$, $P_{0,j}$ and $P_{0,k}$ respectively. At least one of i, j, k is distinct from the other two, and in that case the overall term will be block off-diagonal with respect to the corresponding projector. Hence the overall term provides no contribution to the P_0 space unless $i = j = k$.

In conclusion, we have shown that the only contributions to $(\mathbb{1} \otimes P_0) V^{(3)} (\mathbb{1} \otimes P_0)$ are those which involve no cross-gadget terms. From the assumption that each $h'_j(x)$ is a

gadget for h_j, we are therefore guaranteed that

$$(\mathbb{1} \otimes P_0)V^{(3)}(\mathbb{1} \otimes P_0) = \Big(\sum_{j \in [n]} h_j \Big) \otimes P_0 \ . \tag{4.435}$$

■

Locality of gadgets

Finally, we prove the following simple result, establishing that a strictly k'-local gadget Hamiltonian can only yield a simulated Hamiltonian of locality $\mathcal{O}(dk')$ at order d in its perturbative expansion.

Lemma 4.6.15 (Gadget locality)**.** *Let $H'(x)$ be a gadet of degree d for H. Suppose H is k-local, and H' is k'-local. Then*

$$d \geq \frac{k-1}{k'-1} \ . \tag{4.436}$$

Proof of Lemma 4.6.15. We will show by induction that $V^{(q)}$ is $[q(k'-1)+1]$-local. Note that, by the construction of $T^{(q)}$ in Eq. (4.175) and using Lemma 4.4.1(II), this implies that $T^{(q)}$ is also $[q(k'-1)+1]$-local.

For $q = 1$, we have $V^{(1)} = H^{(1)}$ which is indeed k'-local by assumption. For $q \geq 2$, the result follows by inspection of Eq. (4.161), using the inductive hypothesis for $T^{(q')}$, $q' < q$: each commutator of the form $[T^{(q_1)}, \dots [T^{(q_r)}, H^{(0)}] \dots]$ is at most l-local, where

$$l = \sum_{j=1}^{r} \big(q_j(k'-1)+1 \big) - r + 1 = q(k'-1)+1 \ , \tag{4.437}$$

using that $H^{(0)}$ is 1-local and $\sum_j q_j = q$. Moreover, for $\alpha \geq 1$ each commutator of the form $[T^{(q_1)}, \dots [T^{(q_r)}, H^{(\alpha)}] \dots]$ is at most l'-local, where

$$l' = \sum_{j=1}^{r} \big(q_j(k'-1)+1 \big) - r + k' \leq q(k'-1)+1 \ , \tag{4.438}$$

using that $H^{(\alpha)}$ is at most k'-local and $\sum_j q_j \leq q-1$ for terms of this form. With this in hand, we see that $V^{(d)}$ — and hence H — is $[d(k'-1)+1]$-local. But then by assumption we have

$$k \leq d(k'-1)+1 \ , \tag{4.439}$$

from which the result follows. ■

Quasi-locality of observables on the effective space

Below we prove Lemma 4.5.3 which, informally, ensures that local measurements on the simulator system correspond to measurement of a quasi-local observable on the effective system $O_{A,\mathsf{eff}}(x)$ which can be truncated at finite radius up to small error.

Lemma 4.6.16 (Restatement of Lemma 4.5.3)**.** *Let $O_{A,\mathsf{eff}}(x)$ be the observable on Γ_{eff} defined in Eq.* (4.253)*, and assume $x \leq 1/(2\theta)$. Then, for any $r \geq 0$, there exists an observable $O^{[r]}_{A,\mathsf{eff}}(x)$ on Γ_{eff}, such that $O^{[r]}_{A,\mathsf{eff}}(x)$ is supported on $B_r(A)$, we have*

$$\left\| O_{A,\mathsf{eff}}(x) - O^{[r]}_{A,\mathsf{eff}}(x) \right\| \leq a e^{-bg(r)} \ , \tag{4.440}$$

for some constants $a, b > 0$. Moreover, $O^{[r]}_{A,\mathsf{eff}}(x)$ extends to a complex function $O^{[r]}_{A,\mathsf{eff}}(z)$ which is analytic on the disc $|z| \leq 1/(2\theta)$, and which is bounded as

$$\sup_{|z| \leq 1/(2\theta)} \| O^{[r]}_{A,\mathsf{eff}}(z) \| \leq a' e^{b' r^D} \ , \tag{4.441}$$

for constants $a', b' > 0$.

Proof of Lemma 4.6.16. We define $S(x)$ to be the Hamiltonian

$$S(x) := \frac{T(x)}{ix} \ . \tag{4.442}$$

By the construction of $T(x)$, we are guaranteed that $S(x)$ is an analytic function of x, and by Theorem 4.4.9(I) we know that

$$\| S(x) \|_{F_g} \leq \theta/8 \ . \tag{4.443}$$

Writing $S(x)$ as a sum of local terms $S(x) = \sum_{A' \subseteq \Gamma} S_{A'}(x)$, we define $S^{[r]}(x) := \sum_{A' \subseteq B_r(A)} S_{A'}(x)$ to be the restriction of $S(x)$ which only contains terms supported within a radius r of A. For $0 \leq t \leq x$, we write

$$O_A(x;t) := e^{itS(x)} O_A e^{-itS(x)} \ , \tag{4.444}$$

$$O^{[r]}_A(x;t) := e^{itS^{[r]}(x)} O_A e^{-itS^{[r]}(x)} \ . \tag{4.445}$$

We will now use the locality of $S(x)$, with Lieb-Robinson bounds, to show that $O^{[r]}_A(x;t)$ is a good approximation for $O_A(x;t)$ up to an error which shrinks exponentially with r. Define $\epsilon(x;t) := O_A(x;t) - O^{[r]}_A(x;t)$. This satisfies $\epsilon(x;0) = 0$, and

$$\frac{\partial}{\partial t} \epsilon(x;t) = i[S(x), O_A(x;t)] - i[S^{[r]}(x), O^{[r]}_A(x;t)] \tag{4.446}$$

$$= i[S(x) - S^{[r]}(x), O_A(x;t)] + i[S^{[r]}(x), \epsilon(x;t)] \ . \tag{4.447}$$

This differential equation can be solved to give

$$\epsilon(x;t) = ie^{itS^{[r]}(x)} \left(\int_0^t \mathrm{d}\tau e^{-i\tau S^{[r]}(x)} [S(x) - S^{[r]}(x), O_A(x;\tau)] e^{i\tau S^{[r]}(x)} \right) e^{-itS^{[r]}(x)} \ . \tag{4.448}$$

Using the triangle inequality we can thus bound

$$\|\epsilon(x;t)\| \leq \int_0^t \mathrm{d}\tau \left\| [S(x) - S^{[r]}(x), O_A(x;\tau)] \right\| \ . \tag{4.449}$$

To bound the integrand, we rewrite $S(x) - S^{[r]}(x)$ into a sum of local terms which do not act exclusively on A and apply the triangle inequality, as follows:

$$\left\| [S(x) - S^{[r]}(x), O_A(x;\tau)] \right\| \leq \sum_{\substack{A' \subseteq \Gamma \\ A' \not\subseteq B_r(A)}} \| [S_{A'}(x), O_A(x;\tau)] \| \tag{4.450}$$

$$\leq \sum_{A' \subseteq \Gamma \setminus B_{r/2}(A)} \| [S_{A'}(x), O_A(x;\tau)] \| + \sum_{\substack{A' \subseteq \Gamma \\ A' \cap B_{r/2}(A) \neq \emptyset \\ A' \cap (\Gamma \setminus B_r(A)) \neq \emptyset}} \| [S_{A'}(x), O_A(x;\tau)] \| \ . \tag{4.451}$$

In the second line, we have split the sum into two parts: one containing the terms in $S(x)$ whose support lies a distance at least $r/2$ from A; and one containing the terms in $S(x)$ whose support contains at least one element within distance $r/2$ of A, and at least one element separated by distance r from A. The former will be small due to Lieb-Robinson bounds localising the support of $O_A(x;\tau)$, whilst the latter will be small due to the locality of $S(x)$. In particular, we have:

$$\sum_{A' \subseteq \Gamma \setminus B_{r/2}(A)} \| [S_{A'}(x), O_A(x;\tau)] \| \leq \sum_{r' > r/2} \sum_{\substack{A' \subseteq \Gamma \\ \mathrm{dist}(A', A) = r'}} \| [S_{A'}(x), O_A(x;\tau)] \| \tag{4.452}$$

$$\leq \sum_{r' > r/2} \sum_{\substack{A' \subseteq \Gamma \\ \mathrm{dist}(A', A) = r'}} c \|S'_A(x)\| \|O_A\| |A| \left(e^{\nu \|S(x)\|_{F_g} \tau} - 1 \right) e^{-g(r')} \tag{4.453}$$

$$= c\|O_A\||A| \left(e^{\nu \|S(x)\|_{F_g} \tau} - 1 \right) \sum_{r' > r/2} e^{-g(r')} \sum_{\substack{A' \subseteq \Gamma \\ \mathrm{dist}(A', A) = r'}} \|S_{A'}(x)\| \ , \tag{4.454}$$

where in the second line we have applied Lemma 4.2.14, introducing the constants $c, \nu > 0$. The sums can then be bounded by

$$\sum_{r'>r/2} e^{-g(r')} \sum_{\substack{A'\subseteq\Gamma \\ \mathrm{dist}(A,A')=r'}} \|S_{A'}(x)\| \leq \sum_{r'>r/2} e^{-g(r')} \sum_{i\in B_{r'}(A)} \sum_{\substack{A'\subseteq\Gamma \\ i\in A'}} \|S_{A'}(x)\| \tag{4.455}$$

$$\leq \sum_{r'>r/2} e^{-g(r')} |B_{r'}(A)| \|F_g\| \|S(x)\|_{F_g} \tag{4.456}$$

$$\leq k_D \theta\gamma\Delta^{-1} \|F_g\| \sum_{r'>r/2} (r')^D e^{-g(r')} \,, \tag{4.457}$$

where in the last line we used the bounds $\|S(x)\|_{F_g} \leq \theta/8$ (from Corollary 4.4.3) and $|B_{r'}(A)| \leq k_D (r')^D$ (from Eq. (4.24)). Since we assume exponentially decaying interactions (i.e. g is linear), the right-hand side of this expression is exponentially decaying in r.

Meanwhile, the second term in Eq. (4.451) can be bounded by

$$\sum_{\substack{A'\subseteq\Gamma \\ A'\cap B_{r/2}(A)\neq\emptyset \\ A'\cap(\Gamma\setminus B_r(A))\neq\emptyset}} \|[S_{A'}(x), O_A(x;\tau)]\| \leq 2\|O_A\| \sum_{\substack{A'\subseteq\Gamma \\ A'\cap B_{r/2}(A)\neq\emptyset \\ A'\cap(\Gamma\setminus B_r(A))\neq\emptyset}} \|S_{A'}(x)\| \tag{4.458}$$

$$\leq 2\|O_A\| \sum_{i\in B_{r/2}(A)} \sum_{j\in\Gamma\setminus B_r(A)} \sum_{\substack{A'\subseteq\Gamma \\ i,j\in A'}} \|S_{A'}(x)\| \tag{4.459}$$

$$\leq 2\|O_A\| \sum_{i\in B_{r/2}(A)} \sum_{j\in\Gamma\setminus B_r(A)} F_g(\mathrm{dist}(i,j)) \|S(x)\|_{F_g} \tag{4.460}$$

$$\leq \frac{\|O_A\|\theta}{4} \sum_{i\in B_{r/2}(A)} \sum_{r'>r} \sum_{\substack{j\in\Gamma \\ \mathrm{dist}(A,j)=r'}} F_g(r') \tag{4.461}$$

$$\leq \frac{\|O_A\|\theta F(0)}{4} \sum_{i\in B_{r/2}(A)} \sum_{r'>r} |B_{r'}(A)| e^{-g(r')} \tag{4.462}$$

$$\leq \frac{\|O_A\|\theta F(0) |B_{r/2}(A)|^2}{4} \sum_{r'>r} e^{-g(r')} \tag{4.463}$$

$$\leq \frac{\|O_A\|\theta F(0) |A|^2 k_D^2 (r/2)^D}{4} \sum_{r'>r} (r')^D e^{-g(r')} \,. \tag{4.464}$$

Once again, using the assumption that g is linear, we see that the right-hand side of this expression decays exponentially with r. We can therefore conclude that the bound

Eq. (4.451) decays exponentially in r, and thus

$$\left\| e^{T(x)} O_A e^{-T(x)} - O_A^{[r]}(x;x) \right\| = \|\epsilon(x;x)\| \tag{4.465}$$

$$\leq x \sup_{\tau \in [0,x]} \left\| [S(x) - S^{[r]}(x), O_A(x;\tau)] \right\| \tag{4.466}$$

$$\leq a e^{-bg(r)} \ , \tag{4.467}$$

for some constants $a, b > 0$. Here we have used that $x \leq 1/(2\theta)$ is upper bounded by a constant, and treat $\|O_A\|$ and $|A|$ as constants.

We define the observable $O_{A,\mathsf{eff}}^{[r]}(x)$ on Γ_{eff} by

$$O_{A,\mathsf{eff}}^{[r]}(x) := (\mathbb{1}_{\mathsf{eff}} \otimes \langle \mathbf{0}_{\mathsf{anc}}|) e^{ixS^{[r]}(x)} O_A e^{-ixS^{[r]}(x)} (\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle) \ . \tag{4.468}$$

By definition, $O_{A,\mathsf{eff}}^{[r]}(x)$ has support contained in $B_r(A)$, and we have established that

$$\left\| O_{A,\mathsf{eff}}(x) - O_{A,\mathsf{eff}}^{[r]}(x) \right\| \leq a e^{-bg(r)} \ , \tag{4.469}$$

for $0 \leq x \leq 1/(2\theta)$. Furthermore, since $S^{[r]}(x)$ is constructed as a power series in x, we are guaranteed that $O_{A,\mathsf{eff}}^{[r]}(x)$ extends to an analytic function for $|z| \leq 1/(2\theta)$, bounded by

$$\|O_{A,\mathsf{eff}}^{[r]}(z)\| \leq \|O_A\| \exp\left(2|z| \|S^{[r]}(z)\| \right) \ , \tag{4.470}$$

where we have

$$\|S^{[r]}(z)\| \leq \sum_{A' \subseteq B_r(A)} \|S_{A'}(z)\| \tag{4.471}$$

$$\leq \sum_{i \in B_r(A)} \sum_{\substack{A' \subseteq \Gamma \\ i \in A'}} \|S_{A'}(z)\| \tag{4.472}$$

$$\leq |B_r(A)| \|S(z)\|_{F_g} \|F_g\| \tag{4.473}$$

$$\leq \frac{k_D r^D \theta \|F_g\|}{8} \ , \tag{4.474}$$

so

$$\sup_{|z| \leq 1/(2\theta)} \|O_{A,\mathsf{eff}}^{[r]}(z)\| \leq \|O_A\| \exp\left(\frac{k_D r^D \|F_g\|}{4} \right) \tag{4.475}$$

$$= a' e^{b' r^D} \ , \tag{4.476}$$

for $a', b' > 0$ constants, as required. ∎

4.6.5 Extrapolating extensive quantities

Extrapolation of analytic Hamiltonians

In this section, we will extend Theorem 4.3.2 to the case of extensive observables, i.e. of the form $O = \sum_{A \subseteq \Gamma} O_A$, where O typically has support across the entire system Γ. The only requirement on O is that its interactions decay exponentially over long distances — in other words, it is bounded in the $\|\cdot\|_{F_g}$-norm like the Hamiltonian H.

Theorem 4.6.17. *Let $\{H(x)\}_{x \in \mathcal{X}}$ be a family of Hamiltonians on Γ, where $|\Gamma| = n$. Let O be a (possibly extensive) observable on Γ such that $\|O\|_{F_g} = \mathcal{O}(1)$ (where F_g is the same function describing the exponential decay of interactions in $H(x)$). Then the following holds:*

(I) For constant $\beta > 0$, assume the Gibbs states $\rho_\beta(x)$ satisfy Assumption 4.3.1(I). Then the function $f_\beta(x) := \mathrm{tr}[O\rho_\beta(x)]$ has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = \mathcal{O}(n) \ , \quad x_* \geq R = 1/\,\mathcal{O}(\log^D(n\delta^{-1})) \ . \tag{4.477}$$

(II) Assume that $\{H(x)\}_{x \in \mathcal{S}}$ satisfies Assumption 4.3.1(II). Then the function $f_{\mathsf{ground}}(x) := \langle\psi_0(x)|O|\psi_0(x)\rangle$ has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = \mathcal{O}(n) \ , \quad x_{\min} \geq R = 1/\,\mathcal{O}(\log^{D+1}(n\delta^{-1})) \ . \tag{4.478}$$

(Proof on page 181.)

Concretely, this yields the following scalings for performing Richardson extrapolation:

Corollary 4.6.18 (Extrapolating extensive observables within phases of matter)**.** *The value of $f_\beta(0)$ (respectively $f_{\mathsf{ground}}(0)$) can be calculated up to any desired accuracy $\epsilon > 0$, given the values of $f_\beta(x_k)$ (respectively $f_{\mathsf{ground}}(x_k)$) at m Chebyshev sample points $\{x_k\}_{k=1}^m$, where each x_k is bounded above zero by $x_{\min} := \min_k x_k$, where:*

(I) For f_β,

$$m = \mathcal{O}(\log(n\epsilon^{-1})) \ , \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{D+2}(n\epsilon^{-1}))} \ . \tag{4.479}$$

(II) For f_{ground},

$$m = \mathcal{O}(\log(n\epsilon^{-1})) \ , \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{D+3}(n\epsilon^{-1}))} \ . \tag{4.480}$$

The result still holds using noisy estimates of $f_\beta(x_k)$ (respectively $f_{\mathsf{ground}}(x_k)$ each with additive error $\delta = \Theta(\epsilon/\log\log(\epsilon^{-1}))$.

Proof of Corollary 4.6.18. Corollary 4.2.4 ensures we can extrapolate to within error $\epsilon > 0$, given a (δ, M, R)-analytic approximation and m Chebyshev samples, as long as

$$\epsilon = (\delta + 2^{-m} M)\, \mathcal{O}(\log m) \ . \tag{4.481}$$

This time, we have $M = \mathcal{O}(n)$, so for the right-hand side to be sufficiently small we can choose $m = \mathcal{O}(\log(n\epsilon^{-1}))$ and $\delta = \epsilon / \log\log(n\epsilon^{-1})$. This leads to $R = 1/\,\mathcal{O}(\log^D(n\epsilon^{-1}))$ and $R = 1/\,\mathcal{O}(\log^{D+1}(n\epsilon^{-1}))$ for the Gibbs and ground state cases respectively. Using that $x_{\min} \sim R/m^2$ we arrive at the stated result. ∎

Ultimately, our approach will just be to split O into a sum of $n = |\Gamma|$ observables each localised around a single site, show that each of these has an analytic approximation using Theorem 4.3.2, and conclude by summing the analytic approximations. There is a small subtlety to applying this approach: the individual observables in the summand will in fact be quasi-local rather than strictly local as in the statement of Theorem 4.3.2. To this end, we prove the following lemma (which will also be useful in Section 4.5), generalising Lemma 4.6.4.

Lemma 4.6.19 (Restatement of Lemma 4.5.4)**.** *Let H be a Hamiltonian with bounded $\|\cdot\|_{F_g}$-norm for g linear, and let O be an observable with $\|O\| = \mathcal{O}(1)$ localised around $i \in \Gamma$ in the following sense: for every $r \geq 0$, there exists an observable $O^{[r]}$ with support contained within $B_r(\{i\})$ such that*

$$\|O - O^{[r]}\| \leq a e^{-bg(r)} \ , \tag{4.482}$$

for some constants $a, b > 0$. Let $\Phi_H(O)$ and $\Psi_H(O)$ be the quantum belief propagation and spectral flow operators as defined in Eqs. (4.29) and (4.35). Then there exist constants $a_1', a_2', b_1', b_2' > 0$ such that, for every $r \geq 0$, there exist operators $\Phi_H^{[r]}(O)$ and $\Psi_H^{[r]}(O)$ with support contained within $B_r(\{i\})$ such that

$$\|\Phi_H(O) - \Phi_H^{[r]}(O)\| \leq a_1' \|O\| e^{-b_1' g(r)} \ , \tag{4.483}$$

$$\|\Psi_H(O) - \Psi_H^{[r]}(O)\| \leq a_2' \|O\| e^{-b_2' g(r)/\log^2 g(r)} \ . \tag{4.484}$$

Proof of Lemma 4.6.19. Fix $r \geq 0$, and let $r' = \lfloor r/2 \rfloor$. By assumption, the observable $O^{[r']}$ has support contained within $B_{r'}(\{i\})$, and

$$\|O - O^{[r']}\| \leq a \|O\| e^{-bg(r')} \ . \tag{4.485}$$

Abusing notation slightly, we define $\Phi_H^{[r']}(O^{[r']})$ and $\Psi_H^{[r']}(O^{[r']})$ as prescribed by Lemma 4.2.15, which have support contained within $B_{r'}(B_{r'}(\{i\})) \subseteq B_r(\{i\})$, and where

$$\|\Phi_H(O^{[r']}) - \Phi_H^{[r']}(O^{[r']})\| \leq a_1 |B_{r'}(\{i\})| \|O^{[r']}\| e^{-b_1 g(r')} \ , \tag{4.486}$$

$$\|\Psi_H(O^{[r']}) - \Psi_H^{[r']}(O^{[r']})\| \leq a_2 |B_{r'}(\{i\})| \|O^{[r']}\| e^{-b_2 g(r')/\log^2 g(r')} \ , \tag{4.487}$$

where $a_1, a_2, b_1, b_2 > 0$ are the constants from Eqs. (4.58)-(4.59). By the triangle inequality, we can combine this inequality with Eq. (4.485) to give

$$\|\Phi_H^{[r']}(O^{[r']}) - \Phi_H(O)\| \leq \|\Phi_H(O^{[r']}) - \Phi_H^{[r']}(O^{[r']})\| + \|O - O^{[r']}\| \tag{4.488}$$

$$\leq \left(a_1 |B_{r'}(\{i\})| \|O^{[r']}\| + a\|O\|\right) e^{-b_1 g(r')} \ . \tag{4.489}$$

Since g is linear by assumption, the exponential decay dominates the bracketed term, which grows only polynomially with r', and the result follows by setting $\Phi_H^{[r]}(O) := \Phi_H^{[r']}(O^{[r']})$. An analogous argument applies to $\Psi_H(O)$. ∎

Proof of Theorem 4.6.17. For every $i \in \Gamma$, define the observable $O_{[i]}$ by

$$O_{[i]} := \sum_{\substack{A \subseteq \Gamma \\ i \in A}} \frac{1}{|A|} O_A \ . \tag{4.490}$$

Note that by construction, $O = \sum_{i \in \Gamma} O_{[i]}$. Also, for any $r \geq 0$, we can truncate $O_{[i]}$ to $B_r(\{i\})$ by setting

$$O_{[i]}^{[r]} := \sum_{\substack{A \subseteq B_r(\{i\}) \\ i \in A}} \frac{1}{|A|} O_A \ . \tag{4.491}$$

Then we can bound

$$\|O_{[i]} - O_{[i]}^{[r]}\| \leq \sum_{j \in \Gamma \setminus B_r(i)} \sum_{\substack{A \subseteq \Gamma \\ i,j \in A}} \frac{1}{|A|} \|O_A\| \tag{4.492}$$

$$\leq \sum_{r' > r} \sum_{\substack{j \in \Gamma \\ \mathrm{dist}(i,j) = r'}} \sum_{\substack{A \subseteq \Gamma \\ i,j \in A}} \|O_A\| \tag{4.493}$$

$$\leq \sum_{r' > r} \sum_{\substack{j \in \Gamma \\ \mathrm{dist}(i,j) = r'}} \|O\|_{F_g} F_g(r') \tag{4.494}$$

$$\leq \|O\|_{F_g} \sum_{r' > r} |B_{r'}(i)| F_g(r') \tag{4.495}$$

$$\leq \|O\|_{F_g} F(0) k_D \sum_{r > r'} (r')^D e^{-g(r')} \ . \tag{4.496}$$

Since g is linear, the right-hand side of this expression is exponentially decaying in r and in particular $O_{[i]}$ satisfies the assumptions of Lemma 4.6.19. Now we consider the Gibbs state and ground state cases separately:

(I) For every $i \in \Gamma$, we define the function $f_{\beta,i}(s) := \mathrm{tr}[O_{[x]} \rho_\beta(s)]$. Note that $f_\beta(s) = \sum_{i \in \Gamma} f_{\beta,i}(s)$. We may now apply Theorem 4.3.2(I) to this function —

using Lemma 4.6.19 in place of Lemma 4.6.4, to account for the fact that $O_{[i]}$ is quasi-local rather than strictly local — to deduce that for any $\delta > 0$ there exists a $(\delta/n, \tilde{M}, \tilde{R})$-analytic approximation for $f_{\beta,i}$, where

$$\tilde{M} = \mathcal{O}(1) \ , \quad \tilde{R} = 1/\mathcal{O}(\log^D(n\delta^{-1})) \ . \tag{4.497}$$

We denote this analytic approximation $\tilde{f}_{\beta,i}(z)$, and define the complex function $\tilde{f}_\beta(z)$ as the sum

$$\tilde{f}_\beta(z) := \sum_{i\in\Gamma} \tilde{f}_{\beta,i}(z) \ . \tag{4.498}$$

It follows immediately that $\tilde{f}_\beta$ is a (δ, M, R)-analytic approximation for f_β, where

$$M = n\tilde{M} = \mathcal{O}(n) \ , \quad R = \tilde{R} = 1/\mathcal{O}(\log^D(n\delta^{-1})) \ , \tag{4.499}$$

as required.

(II) Similarly, for every $i \in \Gamma$ we define the function $f_{\mathsf{ground},i}(s) := \langle\psi_0(s)|O_{[i]}|\psi_0(s)\rangle$, and apply Theorem 4.3.2(II) — with Lemma 4.6.19 in place of Lemma 4.6.4 — to construct a $(\delta/n, \tilde{M}, \tilde{R})$-analytic approximation $\tilde{f}_{\mathsf{ground},i}(z)$ for each $f_{\mathsf{ground},i}$, where

$$\tilde{M} = \mathcal{O}(1) \ , \quad \tilde{R} = 1/\mathcal{O}(\log^{D+1}(n\delta^{-1})) \ . \tag{4.500}$$

Summing these gives

$$\tilde{f}_{\mathsf{ground}}(z) := \sum_{i\in\Gamma} \tilde{f}_{\mathsf{ground},i}(z) \ , \tag{4.501}$$

which is a (δ, M, R')-approximation for f_{ground}, where

$$M = n\tilde{M} = \mathcal{O}(n) \ , \quad R = \tilde{R} = 1/\mathcal{O}(\log^{D+1}(n\delta^{-1})) \ , \tag{4.502}$$

as required.

■

Extrapolation of simulator Hamiltonians

In this section, we extend Theorem 4.5.1 to the case of an extensive observable O.

Theorem 4.6.20. *Let $H'(x)$ be a family of gadget Hamiltonians on $\mathcal{H}' = \bigotimes_{i\in\Gamma_{\mathsf{eff}}\cup\Gamma_{\mathsf{anc}}} \mathcal{H}_x$ satisfying the conditions of Assumption 4.4.8 with exponentially decaying interactions, and let $H_{\mathrm{tar}} := H_{\mathsf{eff}}(0)$. Let $O = \sum_{A\subseteq\Gamma_{\mathsf{eff}}} O_A$ be an observable on Γ_{eff} such that $\|O\|_{F_g} = \mathcal{O}(1)$ (where F_g is the same function describing the exponential decay of interactions in $H'(x)$). Then the following holds:*

(I) Let $\beta > 0$, and let $\rho_\beta(x)$, $\rho_{\beta,\mathsf{eff}}(x)$, and $\rho_{\beta,\mathrm{tar}}$ be the Gibbs states corresponding to the Hamiltonians $H'(x)$, $H_{\mathsf{eff}}(x)$, and H_{tar} respectively. Assume the family $\rho_{\beta,\mathsf{eff}}(x)$ satisfies Assumption 4.3.1(I) for $x \in [0, x_*]$. Then the function

$$f'_\beta(x) := \begin{cases} \mathrm{tr}[O\rho'_\beta(x)] & \text{for} \quad x \in (0, x_*) \\ \mathrm{tr}[O\rho_{\beta,\mathrm{tar}}] & \text{for} \quad x = 0 \end{cases} , \tag{4.503}$$

has a (δ, M, R)-analytic approximation for $\delta > 0$, where

$$M = n \exp\left(\mathcal{O}(\log^D(n\delta^{-1}))\right) , \quad x_* \geq R = 1/\,\mathcal{O}(\log^D(n\delta^{-1})) . \tag{4.504}$$

(II) Assume the $H_{\mathsf{eff}}(x)$ satisfies Assumption 4.3.1(II) for $x \in [0, x_*]$, and let $|\psi'_0(x)\rangle$, $|\psi_{0,\mathsf{eff}}(x)\rangle$, and $|\psi_{0,\mathrm{tar}}\rangle$ be the ground states corresponding to the Hamiltonians $H'(x)$, $H_{\mathsf{eff}}(x)$, and H_{tar} respectively. Then the function

$$f'_{\mathsf{ground}}(x) := \begin{cases} \langle\psi'_0(x)|O|\psi'_0(x)\rangle & \text{for} \quad x \in (0, x_*) \\ \langle\psi_{0,\mathrm{tar}}|O|\psi_{0,\mathrm{tar}}\rangle & \text{for} \quad x = 0 \end{cases} , \tag{4.505}$$

has a (δ, M, R)-analytic approximation for any $\delta > 0$, where

$$M = n \exp\left(\mathcal{O}(\log^D(n\delta^{-1}))\right) , \quad x_* \geq R = 1/\,\mathcal{O}(\log^{D+1}(n\delta^{-1})) . \tag{4.506}$$

Proof of Theorem 4.6.20. We proceed exactly as in Theorem 4.6.17, and for every $i \in \Gamma_{\mathsf{eff}}$ we define

$$O_{[i]} := \sum_{\substack{A \subseteq \Gamma \\ i \in A}} \frac{1}{|A|} O_A . \tag{4.507}$$

Then $O = \sum_{i \in \Gamma_{\mathsf{eff}}} O_{[i]}$, and by the same argument as Theorem 4.6.17 we are guaranteed that each $O_{[i]}$ can be truncated to any radius r up to exponentially small error; the same thus holds for the corresponding effective observable

$$O_{[i],\mathsf{eff}}(x) := (\mathbb{1}_{\mathsf{eff}} \otimes \langle \mathbf{0}_{\mathsf{anc}}|) e^{T(x)} O_{[i]} e^{-T(x)} (\mathbb{1}_{\mathsf{eff}} \otimes |\mathbf{0}_{\mathsf{anc}}\rangle) , \tag{4.508}$$

as in Eq. (4.253), by Lemma 4.5.3.

(I) For every $i \in \Gamma_{\mathsf{eff}}$, we define $f'_{\beta,i}(x) := \mathrm{tr}[O_{[i]}\rho'_\beta(x)]$. For $\delta > 0$, Theorem 4.5.1(I) then gives a $(\delta/n, \tilde{M}, \tilde{R})$-analytic approximation $\tilde{f}_{\beta,i}(z)$, where

$$\tilde{M} = \exp\left(\mathcal{O}(\log^D(n\delta^{-1}))\right) , \quad \tilde{R} = 1/\,\mathcal{O}(\log^D(n\delta^{-1})) . \tag{4.509}$$

Hence $\tilde{f}_\beta(z) := \sum_i \tilde{f}_{\beta,i}(z)$ is a (δ, M, R)-analytic approximation for f'_β, where

$$M = n\tilde{M} = n \exp\left(\mathcal{O}(\log^D(n\delta^{-1}))\right) , \quad R = \tilde{R} = 1/\,\mathcal{O}(\log^D(n\delta^{-1})) . \tag{4.510}$$

(II) Similarly, for every $i \in \Gamma_{\mathsf{eff}}$, we define $f_{\mathsf{ground},i}(x) := \langle \psi_0'(x)|O_{[i]}|\psi_0'(x)\rangle$. For $\delta > 0$, Theorem 4.5.1(II) then gives a $(\delta/n, \tilde{M}, \tilde{R})$-analytic approximation $\tilde{f}_{\mathsf{ground},i}(z)$, where

$$\tilde{M} = \exp\left(\mathcal{O}(\log^D(n\delta^{-1}))\right) \ , \quad \tilde{R} = 1/\,\mathcal{O}(\log^{D+1}(n\delta^{-1})) \ . \tag{4.511}$$

Hence $\tilde{f}_{\mathsf{ground}}(z) := \sum_i \tilde{f}_{\mathsf{ground},i}(z)$ is a (δ, M, R)-analytic approximation for f'_{ground}, where

$$M = n\tilde{M} = n\exp\left(\mathcal{O}(\log^D(n\delta^{-1}))\right) \ , \quad R = \tilde{R} = 1/\,\mathcal{O}(\log^{D+1}(n\delta^{-1})) \ . \tag{4.512}$$

■

Corollary 4.6.21 (Extrapolating extensive observables with simulator Hamiltonians)**.** *The value of* $\mathrm{tr}[O\rho_{\beta,\mathrm{tar}}]$ *(respectively* $\langle\psi_{0,\mathrm{tar}}|O|\psi_{0,\mathrm{tar}}\rangle$*) can be calculated up to any desired accuracy* $\epsilon > 0$*, given the values of* $f'_\beta(x_k)$ *(respectively* $f_{\mathsf{ground}}(x_k)$*) at* m *Chebyshev sample points* $\{x_k\}_{k=1}^m$*, where each* x_k *is bounded above zero by* $x_{\min} := \min_k x_k$*, where:*

(I) For f'_β*,*

$$m = \mathcal{O}(\log^D(n\epsilon^{-1})), \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{3D}(n\epsilon^{-1}))} \ . \tag{4.513}$$

(II) For f'_{ground}*,*

$$m = \mathcal{O}(\log^D(n\epsilon^{-1})) \ , \quad x_{\min} = \frac{1}{\mathcal{O}(\log^{3D+1}(n\epsilon^{-1}))} \ . \tag{4.514}$$

In particular, this process requires simulator Hamiltonians with interaction strengths scaling as $x_{\min}^{-d} \sim \mathcal{O}(\mathrm{poly}\log(n\epsilon^{-1}))$*. These conclusions also hold using noisy estimates of* $f'_\beta(x_k)$ *(respectively* $f'_{\mathsf{ground}}(x_k)$*), each with additive error* $\delta = \Theta(\epsilon/\log\log(n\epsilon^{-1}))$*.*

Proof of Corollary 4.6.21. We are guaranteed by Corollary 4.2.4 that extrapolation to error $\epsilon > 0$ is possible with m Chebyshev samples from a (δ, M, R)-analytic approximation, assuming that

$$\epsilon = (\delta + 2^{-m}M)\,\mathcal{O}(\log m) \ . \tag{4.515}$$

By Theorem 4.6.20, we have $M = n\exp(\mathcal{O}(\log^D(n\delta^{-1}))$. Taking $\delta \sim \epsilon/\log\log(n\epsilon^{-1})$, this gives

$$M = n\exp(\mathcal{O}(\log^D(n\epsilon^{-1})) = \exp(\mathcal{O}(\log^D(n\epsilon^{-1})) \ . \tag{4.516}$$

$M = n\exp(\mathcal{O}(\log^D(n\epsilon^{-1}))$. Hence it is sufficient to take $m = \mathcal{O}(\log^D(n\epsilon^{-1}))$. Using $x_{\min} \sim R/m^2$ we arrive at the desired result. ■

References

[1] Dylan Harley, Ishaun Datta, Frederik Ravn Klausen, Andreas Bluhm, Daniel Stilck França, Albert H. Werner, and Matthias Christandl. Going beyond gadgets: the importance of scalability for analogue quantum simulators. *Nature Communications*, 15(1):6527, 2024.

[2] Dylan Harley and Matthias Christandl. Analogue quantum simulation with polylogarithmic interaction strengths by extrapolating within phases of matter. *arXiv preprint arXiv:2605.11285*, 2026.

[3] Mihael Erakovic, Freek Witteveen, Dylan Harley, Jakob Günther, Moritz Bensberg, Oinam Romesh Meitei, Minsik Cho, Troy Van Voorhis, Markus Reiher, and Matthias Christandl. High ground state overlap via quantum embedding methods. *PRX Life*, 3(1):013003, 2025.

[4] Dylan Harley, Freek Witteveen, and Daniel Malz. Computational complexity of injective projected entangled pair states. *arXiv preprint arXiv:2509.19963*, 2025.

[5] Dylan Harley and Robert Koenig. Fault-tolerant one-shot entanglement generation with constant-sized quantum devices in the plane. *arXiv preprint arXiv:2604.05870*, 2026.

[6] Richard P. Feynman. *The Character of Physical Law*. MIT Press, Cambridge, Massachusetts, 1965. Chapter 6: "Probability and Uncertainty: The Quantum Mechanical View of Nature".

[7] Albert Einstein, Boris Podolsky, and Nathan Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical review*, 47(10):777, 1935.

[8] John S. Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195, 1964.

[9] John Bardeen, Leon N. Cooper, and John Robert Schrieffer. Theory of superconductivity. *Physical review*, 108(5):1175, 1957.

[10] Michael A. Nielsen and Isaac L Chuang. *Quantum computation and quantum information*. Cambridge university press, 2010.

[11] Yudong Cao, Jonathan Romero, Jonathan P. Olson, Matthias Degroote, Peter D. Johnson, Mária Kieferová, Ian D. Kivlichan, Tim Menke, Borja Peropadre, Nicolas P. D. Sawaya, et al. Quantum chemistry in the age of quantum computing. *Chemical reviews*, 119(19):10856–10915, 2019.

[12] Alberto Baiardi, Matthias Christandl, and Markus Reiher. Quantum computing for molecular biology. *ChemBioChem*, 24(13):e202300120, 2023.

[13] Bela Bauer, Sergey Bravyi, Mario Motta, and Garnet Kin-Lic Chan. Quantum algorithms for quantum chemistry and quantum materials science. *Chemical reviews*, 120(22):12685–12717, 2020.

[14] Alberto Di Meglio, Karl Jansen, Ivano Tavernelli, Constantia Alexandrou, Srinivasan Arunachalam, Christian W Bauer, Kerstin Borras, Stefano Carrazza, Arianna Crippa, Vincent Croft, et al. Quantum computing for high-energy physics: State of the art and challenges. *Prx quantum*, 5(3):037001, 2024.

[15] Yuri I. Manin. *Computable and Uncomputable*. Sovetskoye Radio, Moscow, 1980. In Russian.

[16] Richard P. Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6-7):467–488, 1982.

[17] Peter W. Shor. Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer. *SIAM review*, 41(2):303–332, 1999.

[18] Lov K. Grover. Quantum mechanics helps in searching for a needle in a haystack. *Physical review letters*, 79(2):325, 1997.

[19] Sean Hallgren. Polynomial-time quantum algorithms for Pell's equation and the principal ideal problem. *Journal of the ACM (JACM)*, 54(1):1–19, 2007.

[20] A. Robert Calderbank and Peter W. Shor. Good quantum error-correcting codes exist. *Physical Review A*, 54(2):1098, 1996.

[21] Dorit Aharonov and Michael Ben-Or. Fault-tolerant quantum computation with constant error. In *Proceedings of the twenty-ninth annual ACM symposium on Theory of computing*, pages 176–188, 1997.

[22] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, et al. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019.

[23] Han-Sen Zhong, Hui Wang, Yu-Hao Deng, Ming-Cheng Chen, Li-Chao Peng, Yi-Han Luo, Jian Qin, Dian Wu, Xing Ding, Yi Hu, et al. Quantum computational advantage using photons. *Science*, 370(6523):1460–1463, 2020.

[24] Youngseok Kim, Andrew Eddins, Sajant Anand, Ken Xuan Wei, Ewout Van Den Berg, Sami Rosenblatt, Hasan Nayfeh, Yantao Wu, Michael Zaletel, Kristan Temme, et al. Evidence for the utility of quantum computing before fault tolerance. *Nature*, 618(7965):500–505, 2023.

[25] Dmitry A. Abanin, Rajeev Acharya, Laleh Aghababaie-Beni, Georg Aigeldinger, Ashok Ajoy, Ross Alcaraz, Igor Aleiner, Trond I Andersen, Markus Ansmann, Frank Arute, et al. Observation of constructive interference at the edge of quantum ergodicity. *Nature*, 646:825—-830, 2025.

[26] Dolev Bluvstein, Simon J. Evered, Alexandra A. Geim, Sophie H. Li, Hengyun Zhou, Tom Manovitz, Sepehr Ebadi, Madelyn Cain, Marcin Kalinowski, Dominik Hangleiter, et al. Logical quantum processor based on reconfigurable atom arrays. *Nature*, 626(7997):58–65, 2024.

[27] I. Trond, Markus Ansmann, Frank Arute, Kunal Arya, Abraham Asfaw, Nikita Astrakhantsev, Juan Atalaya, Ryan Babbush, Dave Bacon, Brian Ballard, et al. Quantum error correction below the surface code threshold. *Nature*, 638:27, 2025.

[28] Microsoft Azure Quantum, Morteza Aghaee, Alejandro Alcaraz Ramirez, Zulfi Alam, Rizwan Ali, Mariusz Andrzejczuk, Andrey Antipov, Mikhail Astafev, Amin Barzegar, Bela Bauer, et al. Interferometric single-shot parity measurement in InAs–Al hybrid devices. *Nature*, 638(8051):651–655, 2025.

[29] Harald Putterman, Kyungjoo Noh, Connor T. Hann, Gregory S. MacCabe, Shahriar Aghaeimeibodi, Rishi N. Patel, Menyoung Lee, William M. Jones, Hesam Moradinejad, Roberto Rodriguez, et al. Hardware-efficient quantum error correction via concatenated bosonic qubits. *Nature*, 638(8052):927–934, 2025.

[30] 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.

[31] Andrew J. Daley, Immanuel Bloch, Christian Kokail, Stuart Flannigan, Natalie Pearson, Matthias Troyer, and Peter Zoller. Practical quantum advantage in quantum simulation. *Nature*, 607(7920):667–676, 2022.

[32] John Preskill. Quantum computing in the NISQ era and beyond. *Quantum*, 2:79, 2018.

[33] Chi-Fang Chen, Michael J. Kastoryano, and András Gilyén. An efficient and exact noncommutative quantum Gibbs sampler, 2023.

[34] Tameem Albash and Daniel A. Lidar. Adiabatic quantum computation. *Reviews of Modern Physics*, 90(1):015002, 2018.

[35] Toby S. Cubitt. Dissipative ground state preparation and the dissipative quantum eigensolver. *arXiv preprint arXiv:2303.11962*, 2023.

[36] Yongtao Zhan, Zhiyan Ding, Jakob Huhn, Johnnie Gray, John Preskill, Garnet Kin-Lic Chan, and Lin Lin. Rapid quantum ground state preparation via dissipative dynamics, 2025.

[37] Chi-Fang Chen, Michael J. Kastoryano, Fernando G. S. L. Brandão, and András Gilyén. Quantum thermal state preparation. *arXiv preprint arXiv:2303.18224*, 2023.

[38] Daniel S. Abrams and Seth Lloyd. Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors. *Physical Review Letters*, 83(24):5162, 1999.

[39] Alexei Yu Kitaev, Alexander Shen, and Mikhail N. Vyalyi. *Classical and quantum computation*. American Mathematical Society, 2002.

[40] Peter W. Shor. Scheme for reducing decoherence in quantum computer memory. *Physical review A*, 52(4):R2493, 1995.

[41] Peter W. Shor. Fault-tolerant quantum computation. In *Proceedings of 37th conference on foundations of computer science*, pages 56–65. IEEE, 1996.

[42] Daniel J. Spencer, Shubham P. Jain, Andrew Tanggara, Zeen Sun, Tobias Haug, Derek Khu, and Kishor Bharti. Quantum error correction and fault tolerance: A comprehensive tutorial. *arXiv preprint arXiv:2605.29137*, 2026.

[43] Seth Lloyd. Universal quantum simulators. *Science*, 273(5278):1073–1078, 1996.

[44] 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.

[45] Dominic W. Berry, Andrew M. Childs, Richard Cleve, Robin Kothari, and Rolando D. Somma. Simulating Hamiltonian dynamics with a truncated Taylor series. *Physical Review Letters*, 114(9):090502, 2015.

[46] Guang Hao Low and Isaac L. Chuang. Hamiltonian simulation by qubitization. *Quantum*, 3:163, 2019.

[47] Guang Hao Low, Vadym Kliuchnikov, and Nathan Wiebe. Well-conditioned multiproduct Hamiltonian simulation. *arXiv preprint arXiv:1907.11679*, 2019.

[48] Earl Campbell. A random compiler for fast hamiltonian simulation. *arXiv preprint arXiv:1811.08017*, 2018.

[49] Joonho Lee, Dominic W. Berry, Craig Gidney, William J. Huggins, Jarrod R. McClean, Nathan Wiebe, and Ryan Babbush. Even more efficient quantum computations of chemistry through tensor hypercontraction. *PRX quantum*, 2(3):030305, 2021.

[50] Isaac H. Kim, Ye-Hua Liu, Sam Pallister, William Pol, Sam Roberts, and Eunseok Lee. Fault-tolerant resource estimate for quantum chemical simulations: Case study on Li-ion battery electrolyte molecules. *Physical Review Research*, 4(2):023019, 2022.

[51] J. Ignacio Cirac and Peter Zoller. Goals and opportunities in quantum simulation. *Nature physics*, 8(4):264–266, 2012.

[52] Yiyi Cai, Yu Tong, and John Preskill. Stochastic error cancellation in analog quantum simulation. *arXiv preprint arXiv:2311.14818*, 2023.

[53] Jayant Rao, Jens Eisert, and Tommaso Guaita. Stability of digital and analog quantum simulations under noise. *arXiv preprint arXiv:2510.08467*, 2025.

[54] Elliott H. Lieb and Derek W. Robinson. The finite group velocity of quantum spin systems. *Communications in mathematical physics*, 28(3):251–257, 1972.

[55] Rahul Trivedi, Adrian Franco Rubio, and J. Ignacio Cirac. Quantum advantage and stability to errors in analogue quantum simulators. *Nature Communications*, 15(1):6507, 2024.

[56] Toby S. Cubitt, Ashley Montanaro, and Stephen Piddock. Universal quantum Hamiltonians. *Proceedings of the National Academy of Sciences*, 115(38):9497–9502, 2018.

[57] Julia Kempe, Alexei Kitaev, and Oded Regev. The complexity of the local Hamiltonian problem. *SIAM Journal on Computing*, 35(5):1070–1097, 2006.

[58] Sanjeev Arora and Boaz Barak. *Computational complexity: a modern approach*. Cambridge University Press, 2009.

[59] John Watrous. Quantum computational complexity. *arXiv preprint arXiv:0804.3401*, 2008.

[60] Sevag Gharibian. Guest column: The 7 faces of quantum NP. *ACM SIGACT News*, 54(4):54–91, 2024.

[61] Hirotada Kobayashi, Keiji Matsumoto, and Tomoyuki Yamakami. Quantum Merlin-Arthur proof systems: Are multiple Merlins more helpful to Arthur? In *International Symposium on Algorithms and Computation*, pages 189–198. Springer, 2003.

[62] John Watrous. PSPACE has constant-round quantum interactive proof systems. *Theoretical Computer Science*, 292(3):575–588, 2003.

[63] Rahul Jain, Zhengfeng Ji, Sarvagya Upadhyay, and John Watrous. QIP = PSPACE. *Journal of the ACM (JACM)*, 58(6):1–27, 2011.

[64] Sanjeev Arora, Carsten Lund, Rajeev Motwani, Madhu Sudan, and Mario Szegedy. Proof verification and the hardness of approximation problems. *Journal of the ACM (JACM)*, 45(3):501–555, 1998.

[65] Zhengfeng Ji, Anand Natarajan, Thomas Vidick, John Wright, and Henry Yuen. Mip* = re. *Communications of the ACM*, 64(11):131–138, 2021.

[66] Chris Marriott and John Watrous. Quantum Arthur–Merlin games. *Computational Complexity*, 14(2):122–152, 2005.

[67] Stephen A. Cook. The complexity of theorem-proving procedures. In *Proceedings of the third annual ACM symposium on Theory of computing*, pages 151–158, 1971.

[68] Leonid A. Levin. Universal sequential search problems. *Problems of information transmission*, 9(3):265–266, 1973.

[69] Tobias J. Osborne. Hamiltonian complexity. *Reports on progress in physics*, 75(2):022001, 2012.

[70] Richard P. Feynman. Quantum mechanical computers. *Foundations of Physics*, 16(6):507–531, 1986.

[71] Alastair Kay. Perfect, efficient, state transfer and its application as a constructive tool. *International Journal of Quantum Information*, 8(04):641–676, 2010.

[72] 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.

[73] Andrew M. Childs. Universal computation by quantum walk. *Physical Review Letters*, 102(18):180501, 2009.

[74] Andrew M. Childs, David Gosset, and Zak Webb. Universal computation by multiparticle quantum walk. *Science*, 339(6121):791–794, 2013.

[75] Karl GH Vollbrecht and J. Ignacio Cirac. Quantum simulators, continuous-time automata, and translationally invariant systems. *Physical Review Letters*, 100(1):010501, 2008.

[76] Alastair Kay. Computational power of symmetric Hamiltonians. *Physical Review A—Atomic, Molecular, and Optical Physics*, 78(1):012346, 2008.

[77] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Physical Review Letters*, 91(14):147902, 2003.

[78] Guifré Vidal. Efficient simulation of one-dimensional quantum many-body systems. *Physical Review Letters*, 93(4):040502, 2004.

[79] Norbert Schuch, Michael M. Wolf, Frank Verstraete, and J. Ignacio Cirac. Computational complexity of projected entangled pair states. *Physical review letters*, 98(14):140506, 2007.

[80] Frank Verstraete and J. Ignacio Cirac. Renormalization algorithms for quantum-many body systems in two and higher dimensions. *arXiv preprint cond-mat/0407066*, 2004.

[81] Jacob Jordan, Roman Orús, Guifre Vidal, Frank Verstraete, and J. Ignacio Cirac. Classical simulation of infinite-size quantum lattice systems in two spatial dimensions. *Physical review letters*, 101(25):250602, 2008.

[82] Julia Kempe and Oded Regev. 3-local Hamitonian is QMA-complete. *Quantum Information and Computation*, 3(3):258–264, 2003.

[83] Daniel Nagaj. Local Hamiltonians in quantum computation. *arXiv preprint arXiv:0808.2117*, 2008.

[84] Jacob D. Biamonte and Peter J. Love. Realizable Hamiltonians for universal adiabatic quantum computers. *Physical Review A*, 78(1):012352, 2008.

[85] Norbert Schuch and Frank Verstraete. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nature physics*, 5(10):732–735, 2009.

[86] Toby Cubitt and Ashley Montanaro. Complexity classification of local Hamiltonian problems. *SIAM Journal on Computing*, 45(2):268–316, 2016.

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

[88] Dorit Aharonov, Daniel Gottesman, Sandy Irani, and Julia Kempe. The power of quantum systems on a line. *Communications in mathematical physics*, 287(1):41–65, 2009.

[89] 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.

[90] 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.

[91] Tamara Kohler, Stephen Piddock, Johannes Bausch, and Toby Cubitt. General conditions for universality of quantum Hamiltonians. *PRX Quantum*, 3(1):010308, 2022.

[92] Sergey Bravyi, David P. Divincenzo, Roberto I. Oliveira, and Barbara M. Terhal. The complexity of stoquastic local Hamiltonian problems. *arXiv preprint quant-ph/0606140*, 2006.

[93] Andrew M. Childs, David Gosset, and Zak Webb. The Bose-Hubbard model is QMA-complete. In *International Colloquium on Automata, Languages, and Programming*, pages 308–319. Springer, 2014.

[94] Ainesh Bakshi, Allen Liu, Ankur Moitra, and Ewin Tang. High-temperature Gibbs states are unentangled and efficiently preparable. In *2024 IEEE 65th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1027–1036. IEEE, 2024.

[95] Cambyse Rouzé, Daniel Stilck França, and Álvaro M. Alhambra. Efficient thermalization and universal quantum computing with quantum Gibbs samplers. In *Proceedings of the 57th Annual ACM Symposium on Theory of Computing*, pages 1488–1495, 2025.

[96] Dorit Aharonov, Itai Arad, and Thomas Vidick. Guest column: the quantum PCP conjecture. *Acm sigact news*, 44(2):47–79, 2013.

[97] Bryan O'Gorman, Sandy Irani, James Whitfield, and Bill Fefferman. Electronic structure in a fixed basis is QMA-complete. *arXiv preprint arXiv:2103.08215*, 2021.

[98] Michael A. Nielsen et al. The Fermionic canonical commutation relations and the Jordan-Wigner transform. *School of Physical Sciences, The University of Queensland*, 59:75, 2005.

[99] Sergey B. Bravyi and Alexei Yu Kitaev. Fermionic quantum computation. *Annals of Physics*, 298(1):210–226, 2002.

[100] Charles Derby, Joel Klassen, Johannes Bausch, and Toby Cubitt. Compact fermion to qubit mappings. *Physical Review B*, 104(3):035118, 2021.

[101] Andrew Tranter, Peter J. Love, Florian Mintert, and Peter V. Coveney. A comparison of the Bravyi–Kitaev and Jordan–Wigner transformations for the quantum simulation of quantum chemistry. *Journal of chemical theory and computation*, 14(11):5617–5630, 2018.

[102] Mitchell Chiew and Sergii Strelchuk. Optimal fermion-qubit mappings. *Bulletin of the American Physical Society*, 68, 2023.

[103] Ivan Kassal, Stephen P. Jordan, Peter J. Love, Masoud Mohseni, and Alán Aspuru-Guzik. Polynomial-time quantum algorithm for the simulation of chemical dynamics. *Proceedings of the National Academy of Sciences*, 105(48):18681–18686, 2008.

[104] Matthew Pocrnic, Ignacio Loaiza, Juan Miguel Arrazola, Nathan Wiebe, and Danial Motlagh. Efficient Simulation of Pre-Born-Oppenheimer Dynamics on a Quantum Computer. *arXiv preprint arXiv:2602.11272*, 2026.

[105] Elliot C. Eklund, Arkin Tikku, Patrick Sinnott, William J. Huggins, Guang Hao Low, Dominic W. Berry, and Ivan Kassal. End-to-end simulation of chemical dynamics on a quantum computer. *arXiv preprint arXiv:2603.19007*, 2026.

[106] Dorit Aharonov and Tomer Naveh. Quantum NP — a survey. *arXiv preprint quant-ph/0210077*, 2002.

[107] John Bostanci, Jonas Haferkamp, Chinmay Nirkhe, and Mark Zhandry. Separating QMA from QCMA with a classical oracle. *arXiv preprint arXiv:2511.09551*, 2025.

[108] Goran Lindblad. On the generators of quantum dynamical semigroups. *Communications in Mathematical Physics*, 48:119–130, 1976.

[109] Martin Kliesch, Thomas Barthel, Christian Gogolin, Michael Kastoryano, and Jens Eisert. Dissipative quantum Church-Turing theorem. *Physical review letters*, 107(12):120501, 2011.

[110] Huanchen Zhai, Chenghan Li, Xing Zhang, Zhendong Li, Seunghoon Lee, and Garnet Kin-Lic Chan. Classical solution of the FeMo-cofactor model to chemical accuracy and its implications. *arXiv preprint arXiv:2601.04621*, 2026.

[111] Markus Reiher, Nathan Wiebe, Krysta M. Svore, Dave Wecker, and Matthias Troyer. Elucidating reaction mechanisms on quantum computers. *Proceedings of the national academy of sciences*, 114(29):7555–7560, 2017.

[112] Seunghoon Lee, Joonho Lee, Huanchen Zhai, Yu Tong, Alexander M Dalzell, Ashutosh Kumar, Phillip Helms, Johnnie Gray, Zhi-Hao Cui, Wenyuan Liu, et al. Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry. *Nature communications*, 14(1):1952, 2023.

[113] Sevag Gharibian and François Le Gall. Dequantizing the quantum singular value transformation: hardness and applications to quantum chemistry and the quantum PCP conjecture. In *Proceedings of the 54th annual ACM SIGACT symposium on theory of computing*, pages 19–32, 2022.

[114] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. Quantum algorithms revisited. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 454(1969):339–354, 1998.

[115] Chris Cade, Marten Folkertsma, Sevag Gharibian, Ryu Hayakawa, François Le Gall, Tomoyuki Morimae, and Jordi Weggemans. Improved hardness results for the guided local Hamiltonian problem. *arXiv preprint arXiv:2207.10250*, 2022.

[116] Matthew B. Hastings and Tohru Koma. Spectral gap and exponential decay of correlations. *Communications in mathematical physics*, 265:781–804, 2006.

[117] Matthew B. Hastings. Solving gapped Hamiltonians locally. *Physical Review B—Condensed Matter and Materials Physics*, 73(8):085115, 2006.

[118] Zeph Landau, Umesh Vazirani, and Thomas Vidick. A polynomial time algorithm for the ground state of one-dimensional gapped local hamiltonians. *Nature Physics*, 11(7):566–569, 2015.

[119] Carlos E. González-Guillén and Toby S. Cubitt. History-state Hamiltonians are critical. *arXiv preprint arXiv:1810.06528*, 2018.

[120] Abhinav Deshpande, Alexey V. Gorshkov, and Bill Fefferman. Importance of the spectral gap in estimating ground-state energies. *PRX Quantum*, 3(4):040327, 2022.

[121] Matthew B. Hastings. Quantum belief propagation: An algorithm for thermal quantum systems. *Physical Review B—Condensed Matter and Materials Physics*, 76(20):201102, 2007.

[122] Matthew B. Hastings and Xiao-Gang Wen. Quasiadiabatic continuation of quantum states: The stability of topological ground-state degeneracy and emergent gauge invariance. *Physical Review B—Condensed Matter and Materials Physics*, 72(4):045141, 2005.

[123] Toby S. Cubitt, David Perez-Garcia, and Michael M. Wolf. Undecidability of the spectral gap. *Nature*, 528(7581):207–211, 2015.

[124] James M. Crawford and Larry D. Auton. Experimental results on the crossover point in random 3-SAT. *Artificial intelligence*, 81(1-2):31–57, 1996.

[125] Marc Mézard, Giorgio Parisi, and Riccardo Zecchina. Analytic and algorithmic solution of random satisfiability problems. *Science*, 297(5582):812–815, 2002.

[126] Francisco Barahona. On the computational complexity of Ising spin glass models. *Journal of Physics A: Mathematical and General*, 15(10):3241, 1982.

[127] Sergey Bravyi and Mikhail Vyalyi. Commutative version of the k-local Hamiltonian problem and common eigenspace problem. *arXiv preprint quant-ph/0308021*, 2003.

[128] Dorit Aharonov and Lior Eldar. On the complexity of commuting local Hamiltonians, and tight conditions for topological order in such systems. In *2011 IEEE 52nd Annual Symposium on Foundations of Computer Science*, pages 334–343. IEEE, 2011.

[129] Norbert Schuch. Complexity of commuting Hamiltonians on a square lattice of qubits. *arXiv preprint arXiv:1105.2843*, 2011.

[130] Sergey Bravyi. Efficient algorithm for a quantum analogue of 2-SAT. *arXiv preprint quant-ph/0602108*, 2006.

[131] David Gosset and Daniel Nagaj. Quantum 3-SAT is QMA1-complete. *SIAM Journal on Computing*, 45(3):1080–1128, 2016.

[132] Daniel Nagaj and Shay Mozes. New construction for a QMA complete three-local Hamiltonian. *Journal of mathematical physics*, 48(7), 2007.

[133] Stephen Piddock. Quantum Max-Cut is NP hard to approximate. *arXiv preprint arXiv:2510.07995*, 2025.

[134] Michael R. Garey, David S. Johnson, and Larry Stockmeyer. Some simplified NP-complete problems. In *Proceedings of the sixth annual ACM symposium on Theory of computing*, pages 47–63, 1974.

[135] 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.

[136] Michael H. Freedman and Matthew B. Hastings. Quantum systems on non-k-hyperfinite complexes: A generalization of classical statistical mechanics on expander graphs. *arXiv preprint arXiv:1301.1363*, 2013.

[137] Anurag Anshu, Nikolas P. Breuckmann, and Chinmay Nirkhe. NLTS Hamiltonians from good quantum codes. In *Proceedings of the 55th Annual ACM Symposium on Theory of Computing*, pages 1090–1096, 2023.

[138] Fernando G. S. L. Brandao and Aram W. Harrow. Product-state approximations to quantum ground states. In *Proceedings of the forty-fifth annual ACM symposium on Theory of computing*, pages 871–880, 2013.

[139] 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.

[140] Iulia M. Georgescu, Sahel Ashhab, and Franco Nori. Quantum simulation. *Reviews of Modern Physics*, 86(1):153–185, 2014.

[141] Stephen Piddock and Johannes Bausch. Universal translationally-invariant Hamiltonians. *arXiv preprint arXiv:2001.08050*, 2020.

[142] Stephen Piddock and Ashley Montanaro. Universal qudit Hamiltonians. *Communications in Mathematical Physics*, 382:721–771, 2021.

[143] 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.

[144] Leo Zhou and Dorit Aharonov. Strongly universal Hamiltonian simulators. *arXiv preprint arXiv:2102.02991*, 2021.

[145] Dorit Aharonov and Leo Zhou. Hamiltonian sparsification and gap-simulations. *arXiv preprint arXiv:1804.11084*, 2018.

[146] 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.

[147] Wolfgang Dür, Guifre Vidal, and J. Ignacio Cirac. Three qubits can be entangled in two inequivalent ways. *Physical Review A*, 62(6):062314, 2000.

[148] Michael Walter, Brent Doran, David Gross, and Matthias Christandl. Entanglement polytopes: multiparticle entanglement from single-particle information. *Science*, 340(6137):1205–1208, 2013.

[149] Matthias Christandl, Vladimir Lysikov, Vincent Steffan, Albert H. Werner, and Freek Witteveen. The resource theory of tensor networks. *arXiv preprint arXiv:2307.07394*, 2023.

[150] Isaac H. Kim and Brian Swingle. Robust entanglement renormalization on a noisy quantum computer. *arXiv preprint arXiv:1711.07500*, 2017.

[151] Johannes Borregaard, Matthias Christandl, and Daniel Stilck França. Noise-robust exploration of many-body quantum states on near-term quantum devices. *npj Quantum Information*, 7(1):45, 2021.

[152] 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.

[153] Yudong Cao, Ryan Babbush, Jacob Biamonte, and Sabre Kais. Hamiltonian gadgets with reduced resource requirements. *Physical Review A*, 91(1):012315, 2015.

[154] Yudong Cao and Daniel Nagaj. Perturbative gadgets without strong interactions. *Quantum Information and Computation*, 15(13-14):1197–1222, 2015.

[155] 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.

[156] Philippe Lewalle, Leigh S. Martin, Emmanuel Flurin, Song Zhang, Eliya Blumenthal, Shay Hacohen-Gourgy, Daniel Burgarth, and K. Birgitta Whaley. A multi-qubit quantum gate using the zeno effect. *Quantum*, 7:1100, 2023.

[157] Carter Ball and Thomas D. Cohen. Zeno effect suppression of gauge drift in quantum simulations. *arXiv preprint arXiv:2405.09462*, 2024.

[158] Eliya Blumenthal, Chen Mor, Asaf A. Diringer, Leigh S. Martin, Philippe Lewalle, Daniel Burgarth, K. Birgitta Whaley, and Shay Hacohen-Gourgy. Demonstration of universal control between non-interacting qubits using the quantum zeno effect. *npj Quantum Information*, 8(1):88, 2022.

[159] David P. DiVincenzo. The physical implementation of quantum computation. *Fortschritte der Physik: Progress of Physics*, 48(9-11):771–783, 2000.

[160] A. Yu Kitaev. Quantum computations: algorithms and error correction. *Russian Mathematical Surveys*, 52(6):1191, 1997.

[161] Henrik Wilming and Albert H. Werner. Lieb-Robinson bounds imply locality of interactions. *Physical Review B*, 105(12):125101, 2022.

[162] Harriet Apel and Toby Cubitt. A mathematical framework for quantum Hamiltonian simulation and duality. *arXiv preprint arXiv:2208.11941*, 2022.

[163] 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.

[164] 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.

[165] 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.

[166] Thomas C. Bohdanowicz and Fernando G.S.L. Brandão. Universal Hamiltonians for exponentially long simulation. *arXiv preprint arXiv:1710.02625*, 2017.

[167] Terry Farrelly. A review of quantum cellular automata. *Quantum*, 4:368, 2020.

[168] 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.

[169] 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.

[170] John Watrous. *The Theory of Quantum Information*. Cambridge University Press, 2018.

[171] Stephen Piddock and Ashley Montanaro. The complexity of antiferromagnetic interactions and 2D lattices. *Quantum Information and Computation*, 17(7-8):636–672, 2017.

[172] Sergey Bravyi, David P. DiVincenzo, and Daniel Loss. Schrieffer–Wolff transformation for quantum many-body systems. *Annals of Physics*, 326(10):2793–2826, October 2011.

[173] Stephen P. Jordan and Edward Farhi. Perturbative gadgets at arbitrary orders. *Physical Review A*, 77(6):062329, 2008.

[174] Johannes Bausch. Perturbation gadgets: Arbitrary energy scales from a single strong interaction. *Annales Henri Poincaré*, 21(1):81–114, 2020.

[175] Baidyanath Misra and E. C. George Sudarshan. The Zeno's paradox in quantum theory. *Journal of Mathematical Physics*, 18(4):756–763, 1977.

[176] Rajendra Bhatia. *Matrix analysis*, volume 169. Springer Science and Business Media, 2013.

[177] Chandler Davis and William M. Kahan. Some new bounds on perturbation of subspaces. *Bulletin of the American Mathematical Society*, 75(4):863–868, 1969.

[178] Chandler Davis and William Morton Kahan. The rotation of eigenvectors by a perturbation. III. *SIAM Journal on Numerical Analysis*, 7(1):1–46, 1970.

[179] 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.

[180] Alexander M. Dalzell, Sam McArdle, Mario Berta, Przemyslaw Bienias, Chi-Fang Chen, András Gilyén, Connor T. Hann, Michael J. Kastoryano, Emil T. Khabiboulline, Aleksander Kubica, Grant Salton, Samson Wang, and Fernando G. S. L. Brandão. *Quantum Algorithms: A Survey of Applications and End-to-end Complexities*. Cambridge University Press, April 2025.

[181] Guoming Wang, Daniel Stilck França, Ruizhe Zhang, Shuchen Zhu, and Peter D. Johnson. Quantum algorithm for ground state energy estimation using circuit depth with exponentially improved dependence on precision. *Quantum*, 7:1167, 2023.

[182] Guoming Wang, Daniel Stilck França, Gumaro Rendon, and Peter D. Johnson. Efficient ground-state-energy estimation and certification on early fault-tolerant quantum computers. *Physical Review A*, 111(1):012426, 2025.

[183] Jerome Lloyd and Dmitry A Abanin. Quantum thermal state preparation for near-term quantum processors. *arXiv preprint arXiv:2506.21318*, 2025.

[184] Trond I. Andersen, Nikita Astrakhantsev, Amir H. Karamlou, Julia Berndtsson, Johannes Motruk, Aaron Szasz, Jonathan A. Gross, Alexander Schuckert, Tom Westerhout, Yaxing Zhang, et al. Thermalization and criticality on an analogue–digital quantum simulator. *Nature*, 638(8049):79–85, 2025.

[185] Kristan Temme, Sergey Bravyi, and Jay M. Gambetta. Error mitigation for short-depth quantum circuits. *Physical review letters*, 119(18):180509, 2017.

[186] Ying Li and Simon C. Benjamin. Efficient variational quantum simulator incorporating active error minimization. *Physical Review X*, 7(2):021050, 2017.

[187] Matthias Christandl, Fulvio Gesmundo, Daniel Stilck França, and Albert H. Werner. Optimization at the boundary of the tensor network variety. *Physical Review B*, 103(19):195139, 2021.

[188] James D. Watson and Jacob Watkins. Exponentially reduced circuit depths using trotter error mitigation. *PRX Quantum*, 6(3):030325, 2025.

[189] Josias Langbehn, George Mouloudakis, Emma King, Raphaël Menu, Igor Gornyi, Giovanna Morigi, Yuval Gefen, and Christiane P Koch. Universal cooling of quantum systems via randomized measurements. *arXiv preprint arXiv:2506.11964*, 2025.

[190] Dominik Hahn, S. A. Parameswaran, and Benedikt Placke. Provably efficient quantum thermal state preparation via local driving. *arXiv preprint arXiv:2505.22816*, 2025.

[191] Zhiyan Ding, Yongtao Zhan, John Preskill, and Lin Lin. End-to-end efficient quantum thermal and ground state preparation made simple. *arXiv preprint arXiv:2508.05703*, 2025.

[192] Cristian Tabares, Dominik S. Wild, J. Ignacio Cirac, Peter Zoller, Alejandro González-Tudela, and Daniel González-Cuadra. Estimating ground-state properties in quantum simulators with global control. *arXiv preprint arXiv:2511.04434*, 2025.

[193] Cambyse Rouzé, Daniel Stilck França, Emilio Onorati, and James D. Watson. Efficient learning of ground and thermal states within phases of matter. *Nature Communications*, 15(1):7755, 2024.

[194] Laura Lewis, Hsin-Yuan Huang, Viet T. Tran, Sebastian Lehner, Richard Kueng, and John Preskill. Improved machine learning algorithm for predicting ground state properties. *Nature Communications*, 15(1):895, 2024.

[195] Lewis Fry Richardson. The approximate arithmetical solution by finite differences of physical problems involving differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society of London A*, 210(459-470):307–357, 1911.

[196] Avram Sidi. *Practical extrapolation methods: Theory and applications*, volume 10. Cambridge University Press, 2003.

[197] Aram W. Harrow, Saeed Mehraban, and Mehdi Soleimanifar. Classical algorithms, correlation decay, and complex zeros of partition functions of quantum many-body systems. In *Proceedings of the 52nd Annual ACM SIGACT Symposium on Theory of Computing*, pages 378–386, 2020.

[198] Hsin-Yuan Huang, Richard Kueng, Giacomo Torlai, Victor V. Albert, and John Preskill. Provably efficient machine learning for quantum many-body problems. *Science*, 377(6613):eabk3333, 2022.

[199] Bruno Nachtergaele, Robert Sims, and Amanda Young. Quasi-locality bounds for quantum lattice systems. I. Lieb-Robinson bounds, quasi-local maps, and spectral flow automorphisms. *Journal of Mathematical Physics*, 60(6):061101, 2019.

[200] Álvaro M. Alhambra. Quantum many-body systems in thermal equilibrium. *PRX Quantum*, 4(4):040201, 2023.

[201] Jürg Fröhlich and Daniel Ueltschi. Some properties of correlations of quantum lattice systems in thermal equilibrium. *Journal of Mathematical Physics*, 56(5):053302, 05 2015.

[202] Huzihiro Araki. Gibbs states of a one dimensional quantum lattice. *Communications in Mathematical Physics*, 14(2):120 – 157, 1969.

[203] D. Pérez-García and A. Pérez-Hernández. Locality estimates for complex time evolution in 1D. *Communications in Mathematical Physics*, 399:929 – 970, 2023.

[204] Sergey Bravyi, Matthew B. Hastings, and Spyridon Michalakis. Topological quantum order: stability under local perturbations. *Journal of mathematical physics*, 51(9), 2010.

[205] Sven Bachmann, Spyridon Michalakis, Bruno Nachtergaele, and Robert Sims. Automorphic equivalence within gapped phases of quantum lattice systems. *Communications in Mathematical Physics*, 309(3):835–871, 2012.

[206] Nilanjana Datta, Jürg Fröhlich, Luc Rey-Bellet, and Roberto Fernández. Low-temperature phase diagrams of quantum lattice systems. ii. convergent perturbation expansions and stability in systems with infinite degeneracy. *Helvetica Physica Acta*, 69(5):752–820, 1996.

[207] Emilio Onorati, Cambyse Rouzé, Daniel Stilck França, and James D. Watson. Provably efficient learning of phases of matter via dissipative evolutions. *arXiv preprint arXiv:2311.07506*, 2023.

[208] Andrea Coser and David Pérez-García. Classification of phases for mixed states via fast dissipative evolution. *Quantum*, 3:174, August 2019.

[209] Yingkang Cao, Suying Liu, Haowei Deng, Zihan Xia, Xiaodi Wu, and Yu-Xin Wang. Robust analog quantum simulators by quantum error-detecting codes. *arXiv preprint arXiv:2412.07764*, 2024.

[210] Roland L. Dobrushin and Senya B. Shlosman. Completely analytical interactions: constructive description. *Journal of Statistical Physics*, 46:983–1014, 1987.

[211] Isaac H. Kim. Perturbative analysis of topological entanglement entropy from conditional independence. *Physical Review B—Condensed Matter and Materials Physics*, 86(24):245116, 2012.

[212] Anurag Anshu, Srinivasan Arunachalam, Tomotaka Kuwahara, and Mehdi Soleimanifar. Sample-efficient learning of interacting quantum systems. *Nature Physics*, 17(8):931–935, 2021.

[213] Toby S. Cubitt, Angelo Lucia, Spyridon Michalakis, and David Perez-Garcia. Stability of local quantum dissipative systems. *Communications in Mathematical Physics*, 337:1275–1315, 2015.

[214] Wasin So and Robert C. Thompson. Singular values of matrix exponentials. *Linear and Multilinear Algebra*, 47(3):249–258, 2000.

[215] Elias M. Stein and Rami Shakarchi. *Complex analysis*, volume 2. Princeton University Press, 2010.

[216] Robert Sims and Gunter Stolz. Many-body localization: concepts and simple models. *arXiv preprint arXiv:1312.0577*, 2013.