12-2016

Combinatorial algorithms for perturbation theory and application on quantum computing

Yudong Cao
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By Yudong Cao

Entitled

Combinatorial Algorithms for Perturbation Theory and Application on Quantum Computing

For the degree of Doctor of Philosophy

Is approved by the final examining committee:

Sabre Kais
Chair

Mikhail Atallah
Co-chair

David Gleich

Ahmed Sameh

Samuel Wagstaff

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Sabre Kais

Approved by: William J. Gorman 8/24/2016

Head of the Departmental Graduate Program Date

COMBINATORIAL ALGORITHMS FOR PERTURBATION THEORY AND APPLICATION ON QUANTUM COMPUTING

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Yudong Cao

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

December 2016

Purdue University

West Lafayette, Indiana

Dedicated to my family and my teachers

ACKNOWLEDGMENTS

I would like to first of all thank my advisor Sabre Kais for his consistent and generous support over the past five years of my graduate studies. Also I would like to thank my collaborators, mainly Anargyros Papageorgiou, Joseph Traub, Jiangfeng Du, Felipe Herrera, Ryan Babbush, Jacob Biamonte, Youhan Fang and Daniel Nagaj. During my pursuit of PhD, I have also benefited from the teaching and helpful discussions with the faculties of Computer Science at Purdue, particularly my committee members: Mike Atallah, David Gleich, Ahmed Sameh and Samuel Wagstaff. In various group meetings, conferences, visits and email communications I have also enjoyed helpful discussions with many great people: Sergio Boixo, Vasil Denchev, Andrew Landahl, Daniel Lidar, James Whitfield, Nathan Wiebe and Zoltan Zimboras. I would also like to thank Qatar Foundation and HBKU for their generous accomodations during my visits to Doha. Finally I acknowledge financial support from National Science Foundation Center for Quantum Information and Computation for Chemistry, Award number CHE-1037992.

PREFACE

This dissertation is interdisciplinary between physics and computer science. The presentation assumes knowledge in basic elements of theoretical computer science but none in physics. The mathematics used is mostly linear algebra. There are in total four chapters. Chapter 1 serves to motivate the central problem that the dissertation concerns. The remaining chapters and appendices are original results obtained in the past three years of research. Most contents in the three chapters have been published; but they also include previously unpublished materials and findings that facilitate the discussions.

The main problem concerned in this dissertation is to use perturbation theory for reducing many-body quantum interactions to realistic two-body ones. Many-body interactions are extremely difficult to implement in experimental conditions, while two-body interactions are far more technologically feasible to realize (for instance D-Wave Systems Inc. has manufactured programmable quantum devices based on two-body interactions to a rather impressive scale). Chapter 1 is intended to argue that 1) many-body interactions arise in a wide variety of contexts in quantum computation, 2) quantum complexity theory offers a powerful set of tools, called *perturbative gadgets*, for reducing many-body interactions to two-body ones and 3) these tools have inherent drawbacks from the perspective of physical realization; and it is the purpose of this dissertation to propose methods for overcoming these drawbacks. In addition, I would like to use Chapter 1 as an opportunity to introduce quantum mechanics and provide simple intuitions on why it is difficult to simulate on classical computers, thus motivating the subject of quantum computing and at the same time provide the conceptual machinery necessary for the developments of later chapters.

Chapter 2 improves the existing constructions of perturbative gadgets. There are also new gadgets that were discovered during the research (Sections 2.8 and 2.6),

which could potentially be of interest. Compared with the published the version [1], the chapter also contains an unpublished Section 2.7. Chapter 3, which is published in [2], proposes a new gadget construction that is entirely different from the existing constructions in the sense that it circumvents the need for strong interactions, which is a common downside of the perturbative gadgets. Apart from its experimental implications, the gadget construction in Chapter 3 is also used in an important theoretical development which shows a counterexample to the area law conjecture in condensed matter physics [3].

The title of this dissertation focuses on combinatorics of perturbation theory and mentions quantum computing as an application. Although Chapters 2 and 3 appear to be entirely devoted to perturbative gadgets, whose primary application is in quantum computing and quantum complexity theory, a hidden theme that develops at their core is in fact a continuously deepened understanding of perturbation theory. In Chapter 2 we are able to improve some existing gadget constructions by a careful examination of the perturbation series. In Chapter 3, in order to prove that the perturbation series expansion converges (Section 3.3.2), we adopt combinatorial analyses of perturbation series that are more involved than those in Chapter 2. It is these analyses that uncovered the association between the perturbative expansion and Motzkin walks, paving the way for more general algorithms in Chapter 4. Essentially, improving perturbative gadgets boils down to finding a tighter upper bound to the norm of the perturbation series from a certain order to infinity. Chapters 2 and 3 deal with this task for specific gadget constructions while the algorithms in Chapter 4 deal with far more general settings (Section 4.2.1). In this sense, Chapters 2 and 3 build up to Chapter 4, which is highlighted in the title as the strongest result of the dissertation. Preliminary version of Chapter 4 is available online [4]. However, in the dissertation there are additional rigorous analyses (Section 4.5.2) which provide evidence as to why our algorithms are able to find tight upper bound to the norm of terms at arbitrary order in the perturbation series.

TABLE OF CONTENTS

Page

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

SYMBOLS

$[m]$	Set of integers $\{1, 2, \cdots, m\}$
$\mathbb{C}$	Set of complex numbers; $\mathbb{C}^D$ is for D-dimensional complex vector space; $\mathbb{C}^{m\times n}$ for the space of $m \times n$ complex matrices
$\lvert S\rvert$	The cardinality of a set S
$\mathbf{A}^\dagger$	Conjugate transpose of matrix $\mathbf{A}$
$\mathbf{A}^{\otimes n}$	n-fold tensor product $\underbrace{\mathbf{A} \otimes \mathbf{A} \otimes \cdots \otimes \mathbf{A}}_{n \text{ times}}$
$\lVert\mathbf{A}\rVert_2$	2-norm of a matrix $\mathbf{A}$, which is defined as $\max_{\mathbf{x}} \lVert\mathbf{Ax}\rVert_2/\lVert\mathbf{x}\rVert_2$ where $\lVert\mathbf{x}\rVert_2 = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$
$\lVert\mathbf{A}\rVert_\infty$	∞-norm (infinity norm) of a matrix $\mathbf{A}$, which is defined as $\max_{i\in[m]} \lvert a_{ij}\rvert$ for $\mathbf{A} \in \mathbb{C}^{m\times n}$
$\lvert\psi\rangle$	A *ket*, which describes a quantum state. In our context $\lvert\psi\rangle$ is a unit column vector in $\mathbb{C}^N$
$\langle\psi\rvert$	A *bra*, which in our context is a unit row vector in $\mathbb{C}^N$. Also for the same ψ, the conjugate transpose of $\lvert\psi\rangle$ is $\langle\psi\rvert$
$\lvert 0\rangle$	The '0' state of a qubit, which is equal to the basis vector $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$
$\lvert 1\rangle$	The '1' state of a qubit, which is equal to the basis vector $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$
$\mathbf{X}$	Pauli X operator, $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
$\mathbf{Y}$	Pauli Y operator, $\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$

$\mathbf{Z}$	Pauli Z operator, $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$
$\mathbf{I}$	Identity operator. Its dimension can be inferred from the context
$\mathbf{X}_i$	Pauli X operator acting on the i-th qubit. For n qubits, $\mathbf{X}_i = \mathbf{I}^{\otimes(i-1)} \otimes \mathbf{X} \otimes \mathbf{I}^{\otimes(n-i)}$. The same notation for Pauli Y and Z operators
$\hbar$	Reduced Planck's constant, $1.054\ 571\ 800(13) \times 10^{-34}$ J $\cdot$ s
q_e	Electron charge, $1.602\ 176\ 62 \times 10^{-19}$ coulombs
$\mathbf{H}$	General notation for a Hamiltonian; (perturbation theory) the unperturbed Hamiltonian
$\mathbf{V}$	(perturbation theory) The perturbation
$\widetilde{\mathbf{H}}$	(perturbation theory) The perturbed Hamiltonian $\mathbf{H} + \mathbf{V}$
$\mathbf{G}(z)$	(perturbation theory) Operator-valued resolvent
E_*	An energy cutoff value. Any state with energy lower than E_* belongs to $\mathcal{L}_-$. Otherwise it belongs in $\mathcal{L}_+$
$\mathcal{L}_-$	The low-energy subspace of a Hamiltonian $\mathbf{H}$
$\mathcal{L}_+$	The high-energy subspace of a Hamiltonian $\mathbf{H}$
$\mathbf{\Pi}_-$,$\mathbf{\Pi}_+$	Projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$
$\mathcal{H}$	The Hilbert space of a certain Hamiltonian. $\mathcal{H} = \mathcal{L}_- \oplus \mathcal{L}_+$

ABBREVIATIONS

AQC	Adiabatic Quantum Computing
BPP	Bounded-error probabilistic polynomial time, the class of decision problems solvable by a probabilistic Turing machine in polynomial time with an error probability bounded away from 1/3 for all instances
BQP	Bounded-error Quantum Polynomial time, quantum analogue of the complexity class BPP
MA	Merlin-Arthur, the set of decision problems that can be decided in polynomial time by an Arthur-Merlin protocol
NMR	Nuclear Magnetic Resonance
QA	Quantum Annealing
QMA	Quantum Merlin-Arthur, quantum analogues of the complexity class MA
RS	Rayleigh-Schrödinger perturbation theory (Section 1.3.1)
SAT	Circuit Satisfiability problem

NOMENCLATURE

cellular automaton	A computational model consisting of a network of basic units called *cells* whose states are governed by simple update rules
classical computer	A computational device that stores and manipulates bits. The term is often used in contrast with *quantum computers*
classical Ising model	A physical system whose state is described by a binary string $(s_1, s_2, \cdots, s_n)$ with energy given by $H = \sum_i h_i s_i + \sum_{i<j} J_{ij} s_i s_j$ where $h_i \in \mathbb{R}$ and $J_{ij} \in \mathbb{R}$
eigenstate	A quantum state that is an eigenvector of a Hamiltonian
entangled state	A state that cannot be expressed as a product state
entanglement	The phenomenon that a quantum system can be in an entangled state
gadget	(computational complexity theory) A finite structure which maps a set of constraints from one optimization problem into a constraint of another problem
ground state	An eigenstate of a physical system described by a Hamiltonian $\mathbf{H}$ with the lowest energy. In other words, it is an eigenvector of $\mathbf{H}$ with the smallest eigenvalue

Hamiltonian	A Hermitian matrix describing the interactions in which a quantum system is involved
perturbation theory	A general method for finding an approximate solution to a problem which can be decomposed into an exactly solvable problem plus a small perturbation
Hilbert space	A complex vector space with an inner product
product state	A quantum state of N subsystems which can be written as a tensor product of the states of the individual subsystems. Often used to contrast *entangled state*
quantum algorithm	An algorithm that runs on a quantum computer. Often a quantum algorithm is a sequence of physically realizable operations
quantum computer	A computational device that stores and manipulates qubits. Often synonymous with *universal quantum computer*, which is also capable of realizing arbitrary unitary transformation with arbitrarily small error
quantum state	The state of a quantum system. In our context a quantum state is represented by a unit vector of complex numbers whose 2-norm is unity
quantum Turing machine	An abstract model of quantum computing that generalizes the construction of Turing machine to include quantum mechanics as its working principle [5, 6]
qubit	A unit of quantum information whose states are 2-dimensional complex vectors $\binom{\alpha}{\beta}$ such that $\lvert\alpha\rvert^2 + \lvert\beta\rvert^2 = 1$, $\alpha, \beta \in \mathbb{C}$

resolvent	(complex analysis) An operator $\mathbf{R}$ that captures the spectrum of another operator $\mathbf{A}$ in its structure. Specifically, if we consider the operator as a constant, we have the operator-valued resolvent $\mathbf{R}(z) = (z\mathbf{I} - \mathbf{A})^{-1}$
Schödinger equation	The governing equation for the quantum mechanical behaviours of any closed physical system
symmetric polynomial	A polynomial $f(x_1, x_2, \cdots, x_n)$ whose value is invariant with respect to any permutation of the variables
(physical) system	A collection of physical objects such as atoms and molecules. In a *quantum system* such objects interact according to laws of quantum mechanics
universal gate set	A set of elementary quantum operations from which one could approximate any unitary operation efficiently by forming a sequence of operations from the set

ABSTRACT

Cao, Yudong Ph.D., Purdue University, December 2016. Combinatorial Algorithms for Perturbation Theory and Application on Quantum Computing. Major Professor: Sabre Kais.

This dissertation concerns the problem of simulating arbitrary quantum many-body interactions using realistic two-body interactions. To address this issue, a general class of techniques called perturbative reductions (or perturbative gadgets) is adopted from quantum complexity theory and in this dissertation these techniques are improved for experimental considerations. The idea of perturbative reduction is based on the mathematical machinery of perturbation theory in quantum physics. A central theme of this dissertation is then to analyze the combinatorial structure of the perturbation theory as it is used for perturbative reductions. Specifically, the original contribution of this dissertation is three-fold:

1. Improvement over existing perturbative reductions [7–9] by reducing the resources needed for realizing them [1].

2. Proposal of a new perturbative reduction [2] that circumvents the need for strong interaction in almost all existing constructions, providing a more practical alternative for realizing many-body simulation. Here we go beyond the usual regime of convergence for perturbation theory where the perturbation as a whole is much smaller than the unperturbed section.

3. An efficient algorithm for computing tight upper bounds to perturbation series at arbitrary order [4]. The algorithm deals with a much broader class of physical settings and treats the combinatorial structure of perturbative expansion in much greater generality than the specific analyses for perturbative reductions.

1. INTRODUCTION

1.1 Overview

Quantum computing is an emerging field at the intersection between computer science and physics. The original ideas for quantum computing could be traced back to the 1980s in the works of Richard Feynman [10,11] and Yuri Manin [12]. Feynman's observation was that classical computers[1] take an exponentially long time to simulate quantum many-body systems[2]. If this is an inherent property of quantum mechanics, then one could imagine a computer exploiting quantum mechanical effects may be able to perform such simulations exponentially faster than classical computers.

This idea of a quantum mechanical computer was later formally captured by *quantum Turing machine*, which generalizes the notion of a Turing machine to include quantum mechanics as its working principle [5, 6]. This sets the stage for further development in computational complexity theory that seeks to identify which problems are efficiently solvable using a universal quantum computer and which problems are likely not. A major breakthrough was made by Peter Shor [13, 14] for finding an algorithm that solves integer factorization in polynomial time using quantum computers. Because of the presumed exponential cost of factoring on classical computers, which underlies most of today's crypto systems, the discovery of Shor's algorithm greatly motivated institutions around the world to realize scalable quantum computers. Although in theory there is no obstacle that prevents quantum computers from being built, experimental progress so far has been rather slow. However, theoretical

[1]The term *classical computer* broadly refers to computational devices that store and manipulate bits, as opposed to *quantum computers*, which stores and manipulates *qubits*.

[2]A physical *system* means a collection of physical objects such as atoms and molecules. In a *quantum system* such objects interact according to laws of quantum mechanics

developments born out of the interactions between computer scientists and physicists have been quite fruitful.

In classical computational complexity theory, the complexity class P refers to the set of problems that can be solved efficiently on a classical computer and NP refers to those whose solutions can be checked efficiently on a classical computer. With the formal definition of a quantum Turing machine, the quantum analogues of P and NP could also be introduced. The class of problems that are efficiently solvable on a quantum computer is called *Bounded-error quantum polynomial time*, or BQP [6]. Shor's result on integer factorization [13, 14] could then be interpreted as saying that integer factorization is in BQP. Of greater interest to this dissertation is the complexity class QMA, which is short for *Quantum Merlin-Arthur*. Loosely speaking QMA can be considered as the quantum analogue[3] of NP [15, 16]. Every problem in QMA has the property that given a solution in the form of a quantum state, one could efficiently check the validity of the solution using a quantum computer in polynomial time. If a problem being NP-complete can be regarded as an evidence that it is hard to solve on a classical computer, a problem being QMA-complete can be thought of it being hard to solve even with a quantum computer.

The formal definitions of BQP and QMA are not merely mathematical exercises put forward by complexity theorists, but meaningful characterizations that provide valuable insights on quantum physics and quantum chemistry. As the recent program of Quantum Hamiltonian Complexity [17] has unveiled, many important problems in condensed matter physics [7,8] and quantum chemistry [18–22] are QMA-hard. These QMA-hardness shows for the first time that ultimate limitations in solving some of the difficult open problems involving quantum mechanics come from a fundamental computational complexity of the problem rather than a lack of human ingenuity, in the same sense as a problem being NP-hard shows a fundamental difficulty of solving the problem instead of a lack of smart ideas for devising a polynomial time algorithm.

[3]In a more strict and subtle sense QMA is a closer analogue to MA (short for Merlin-Arthur) and BQP is closer to BPP. Here BPP and MA are probabilistic analogues of P and NP.

Aside from insights on the quantum mechanical problems themselves, the complexity specifications of problems in quantum physics and chemistry also have inspired experimental efforts in engineering quantum systems that manifest the specified computational complexity. For example, finding the lowest energy configuration (the ground state) of an Ising model is NP-complete [23], which means if one were to build a physical device with sufficient degrees of freedom to realize any instance of Ising model, one may use the device as a tool for solving NP-complete problems (though usually without a guarantee on efficiency). Indeed, the Canadian company D-Wave Systems has based its business model on this fact and manufactured a variant of Ising system which can be physically controlled to an impressive scale [24–27]. Although whether the D-Wave quantum devices could produce algorithmic advantage over the best known classical approaches still remains to be determined, the quantum hardware constructed to date represents a crucial step towards the realization of scalable quantum computers.

Cook-Levin theorem [28–30] is a classical result in computational complexity which says that SAT is NP-complete. From a computational perspective, it shows that SAT, a problem with deceptively simple structure, has enough degrees of freedom to efficiently describe arbitrary finite computational processes of a classical computer (a Turing machine). In other words, if we were given an oracle that solves SAT we could use it as a universal Turing machine by feeding it appropriately constructed SAT instances. D-Wave's quantum devices do not qualify as oracles for solving all instances of the Ising model, but they certainly provide a heuristic which exploits physical effects that are previously unavailable to any classical computers. Since Ising model is also NP-complete, in principle one could construct instances of the Ising model to efficiently simulate all polynomial-time algorithms run by a classical computer. A natural question to ask for the next step of this development would be: what kind of physical systems could efficiently simulate all polynomial time algorithms run by a *quantum computer*?

As a first thought for answering this question, if we draw analogy from our discussion about Cook-Levin theorem, our physical system needs to manifest some form of QMA-completeness. In other words it has to be "complex" enough to capture arbitrary process of quantum computation. Indeed there is a quantum version of Cook-Levin theorem [15, 17, 31] that contructs a specific form of quantum system to efficiently simulate any processes of quantum computation. However, the construction initially proposed [31, Sec. 14.4] requires the quantum particles to engage in many-body interactions, which is highly non-trivial to physically realize in a lab. Our current technology is mostly limited to implementing controllable two-body interactions with additional restrictions on the their strengths and geometry (such as a linear chain or a square lattice). It is then an important problem to reduce arbitrary many-body interactions to the type of interactions that are more physically viable. In the past years complexity theorists have worked hard to find the simplest QMA-complete physical systems [32]. Many constructions of QMA-complete Hamiltonians requiring only two-body interactions on a square lattice [8] or restricted types of interactions [33] have been proposed.

A key technique for accomplishing such reduction from many-body to restricted two-body problems is based on the framework of *perturbation theory* in quantum mechanics. The basic idea of perturbation theory is quite simple, and it applies to a broad class of problems in physics and engineering. Suppose we have a problem $\tilde{H}$ that can be partitioned into two parts: 1) a subproblem H whose solution is known and 2) a perturbation V that has some rather irregular structures that render $\tilde{H}$ much harder to solve directly than H. We could start from a solution ψ of H and obtain a series expansion $\psi' = \psi_0 + \psi_1 + \psi_2 + \psi_3 + \cdots$ that approximates the corresponding solution of $\tilde{H}$. Here usually the zeroth order contribution $\psi_0 = \psi$ and the higher order contributions ψ_i takes into account the influence of V on ψ. If the influence of V in the overall problem $\tilde{H}$ is very small, commonly the contribution of each order ψ_i to ψ' gets smaller as i increases such that the infinite series ψ' converges to a finite limit. For many problems in quantum physics the series diverges and there are methods for

handling such cases. This is not the setting of interest in this dissertation; we only deal with convergent series.

Apart from the expansion for the approximate solution to the perturbed problem, using perturbation theory we could also have an expansion for a specific part of the problem $\tilde{H}$ that we are interested in. In our setting a physical system is described using a Hermitian matrix called a *Hamiltonian*. Suppose we have a quantum system described by a Hamiltonian $\widetilde{\mathbf{H}}$. Our $\widetilde{\mathbf{H}}$ is a sum of some diagonal Hamiltonian $\mathbf{H}$, whose eigenvalues and eigenvectors are trivially known, and some perturbation $\mathbf{V}$ with more complicated off-diagonal structures. In quantum physics the eigenvalues of a Hamiltonian are the energy levels of the physical system[4] and often one is more interested in the subspace spanned by eigenvectors (called *eigenstates*) with the lowest few eigenvalues of $\widetilde{\mathbf{H}}$. The projection of $\widetilde{\mathbf{H}}$ into this subspace gives rise to the effective low-energy physics of the system $\widetilde{\mathbf{H}}$. We could obtain an expansion that approximates this low-energy effective Hamiltonian, which takes the form $\mathbf{\Sigma}_{-} = \mathbf{T}_0 + \mathbf{T}_1 + \mathbf{T}_2 + \cdots$. Here $\mathbf{T}_0$ is the projection of $\mathbf{H}$ onto its own low-energy subspace and the higher order terms $\mathbf{T}_i$ correspond to how the perturbation $\mathbf{V}$ influences the low-energy subspace of $\mathbf{H}$ to turn it into the low-energy subspace of $\widetilde{\mathbf{H}}$.

Back to the reduction from many-body to restricted two-body problems. Suppose we have a many-body system $\mathbf{H}_{\text{targ}}$. We then construct a two-body Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ such that the perturbation theory expansion $\mathbf{\Sigma}_{-}$ for the low-energy effective Hamiltonian gives rise to $\mathbf{H}_{\text{targ}}$ in its leading orders and the remaining terms in the infinite sum are errors that can be suppressed, since we always assume convergence. The constructions for $\widetilde{\mathbf{H}}$ are called *perturbative reduction* (or *perturbative gadgets*). In this dissertation, I will describe different gadget constructions for different purposes but the underlying idea is the same: for a *target Hamiltonian* $\mathbf{H}_{\text{targ}}$ of certain form (for example, containing many-body interaction), we could construct a

[4]...thus the name "quantum" physics because the physical objects have only discrete energy levels, or energy *quanta*, rather than the continuous energy manifested in macroscopic objects. For example when a car accelerates from 0 to 100mph its energy undergoes a continuous change, while a hydrogen atom in its ground state takes a specific amount of energy to jump into its first excited state.

gadget Hamiltonian $\widetilde{\mathbf{H}}$ of simpler form (for example, two-body interaction, restricted interaction type or restricted geometry) such that we could find the ground state of $\widetilde{\mathbf{H}}$ if and only if we could do the same for the target Hamiltonian $\mathbf{H}_{\text{targ}}$. Using perturbative reduction, we are able to reduce many-body systems whose ground state is QMA-complete to find to simpler two-body ones that are presumably more experimentally realizable.

Perturbative reductions certainly provide a tempting option for simulating arbitrary many-body interactions using simple two-body ones. However, for practical realization of these gadget constructions, one important issue remains to resolved. In order to guarantee convergence of the perturbation series, the unperturbed part of the gadget constructions, $\mathbf{H}$, often needs to be large. In fact if the original many-body system $\mathbf{H}_{\text{targ}}$ contains n particles, $\mathbf{H}$ often contains terms whose norm need to scale as poly(n) for increasing n, in order to make sure that the error terms[5] in the perturbative expansion $\mathbf{\Sigma}_{-}$ have magnitude below a fixed threshold. The norm of a term in a Hamiltonian often characterizes the strength of interaction it describes. The polynomial scaling of some terms in $\mathbf{H}$ is then impractical because natural physical interactions are always *local* - an atom in a crystal lattice always interacts most strongly with its nearest neighbors and much less so with its next-nearest ones and so on, regardless of how large the lattice actually is. Hence in order to make the perturbation reductions practical, it is critical to reduce the norm of the terms in $\mathbf{H}$ as much as possible while still not exceeding the error threshold in the perturbative expansion.

This dissertation makes progress in dealing with this issue in two ways. First, it improves on existing constructions [7, 8] of perturbative reductions by deriving tighter upper bounds for the norm of the error terms than previously known and also numerically demonstrates that the interaction strengths of the gadget Hamiltonian constructions can be reduced by orders of magnitude while the error is still below the

[5]Recall that in the expansion $\mathbf{\Sigma}_{-}$ for the low-energy effective Hamiltonian of the gadget $\widetilde{\mathbf{H}}$, the leading orders $\mathbf{T}_1 + \cdots + \mathbf{T}_k$ gives $\mathbf{H}_{\text{targ}}$ while the remaining terms $\mathbf{T}_{k+1} + \cdots$ are error terms.

given threshold [1]. Second, it proposes a new gadget construction [2] which entirely avoids the need for strong interactions. Its perturbation theory analysis goes beyond the usual convergent regime where the magnitude of the perturbation as a whole is much smaller than the unperturbed part of the problem. Regardless, we are able to show convergence of the perturbation series using an observation that maps the terms in the perturbation series at any order r to different r-step Motzkin paths. Over the process of developing the dissertation research, this observation has eventually led to the combinatorial algorithms for perturbation theory based on cellular automata [4], which is highlighted as the strongest result of the dissertation.

For convergent perturbation series, it is not hard to obtain an upper bound to the error $\|\mathbf{T}_k+\mathbf{T}_{k+1}+\cdots\|$ after a certain order k because each term $\mathbf{T}_r$ in $\mathbf{\Sigma}_-(z)$ contains a matrix power $\mathbf{A}^k$ for some matrix $\mathbf{A}$ and an exponent $k = O(r)$. Hence it suffices to find an upper bound to $\|\mathbf{A}\|$ and apply geometric series formula to calculate the total error bound $\|\mathbf{A}\|^k + \|\mathbf{A}\|^{k+1} + \cdots$ from the r^{th} order to infinity (see for example [9, Appendix]). However, what is difficult is to find a *tight* upper bound for the error. The geometric series approach almost always fails to obtain a tight bound because it does not take into account the details of matrix multiplication in $\mathbf{A}^k$, while for many quantum many-body systems the matrix $\mathbf{A}$ being exponentiated often admits certain structures that have non-trivial consequences for its own exponentiation. If we consider the geometric series approach as one extreme which completely ignores the structure of matrix product in $\mathbf{A}^k$, the other extreme would be to simply compute $\mathbf{A}^k$ explicitly for each order and evaluate its norm directly. Of course since $\mathbf{\Sigma}_-$ is an infinite series, it is unfeasible to carry out calculation to infinite order. But even if we ignore this issue for now, because quantum mechanics dictates that the dimensions of the matrices involved in $\mathbf{\Sigma}_-$ scale *exponentially* (refer to the discussion in the opening paragraph of this section) with respect to the size of the physical system, the matrix $\mathbf{A}$ itself easily becomes too large to store for even moderately-sized systems. In short, we have one extreme (geometric series) which requires little computation but yields highly inaccurate results and the other which yields entirely accurate results but

requires an infeasible amount of computation. A natural question to ask would be whether it is possible to find a middle ground where a sufficiently tight error bound is obtained without doing an exponential amount of computation.

The set of algorithms proposed in [4] resolve this issue by building on the intuition established in earlier work [2] that each element of the matrix $\mathbf{T}_r$ is a sum of contributions from r-step Motzkin paths of specific properties. The number of Motzkin walks scales exponentially as the size of the physical system, just as quantum mechanics dictates. However, by exploiting the permutational symmetry in the set of Motzkin paths we show that there is a polynomial time algorithm that sums over all of the Motzkin paths that contribute to a specific matrix element of $\mathbf{T}_r$. The basic model of computation that our algorithm assumes is *cellular automata*. A cellular automaton consists of a collection of interconnected cells. Each cell has a *state* which can be a discrete or continuous value. An initial state of the automaton is defined by assigning a state for each cell. The states of the cells then evolves according to some *update rules* which changes the current state of a cell only based on that of the cells that it connects to. During an *evolution* of the automata, the update rules are applied to all cells simultaneously. In the constructions of our cellular automaton, the states of each cell contain parameters for *symmetric polynomials*[6]. These symmetric polynomials provide a compact representation of different Motzkin walks and the update rules for the cells are set up in a way such that r evolutions correspond exactly to summing over all r-step Motzkin walks of the desired kind for $\mathbf{T}_r$. The connection between Motzkin paths and automata has been known from the perspective of symmetric polynomials [34]. Our result [4] can be considered as combining that connection with the tensor product structure of quantum mechanics to yield an efficient and accurate procedure for estimating the error in perturbation theory.

[6]A symmetric function is a function whose value does not change if we permute the variables in any way. For example $f(x, y, z) = f(y, z, x)$ if f is a symmetric function over all three variables. If f is furthermore a polynomial then f is a symmetric polynomial.

1.2 Quantum mechanics

It is hard to define exactly what quantum mechanics is using a concise, self-contained statement. For example, a common definition is that quantum mechanics is "the physics of the very small". However, there are macroscopic objects that also obey the laws of quantum mechanics, such as a block of superconductor at low temperature. Because of the extreme precision with which the predictions of quantum mechanics have been verified over the past century, combined with the seemingly strange physical picture that it suggests, the possible interpretations of quantum mechanics have triggered endlessly fascinating philosophical debates since the beginning of the last century. On the practical side, applications of quantum mechanics on a myriad of physical systems have also led to remarkable advances in science and technology such as the laser, the transistor and magnetic resonance imaging (MRI). Much is also to be said about the mathematical framework of quantum mechanics and its deep roots in group theory. But here we pursue a minimalist introduction to the aspects of quantum mechanics that are relevant to this dissertation. For more comprehensive treatment, one could refer to any of the classic texts on quantum mechanics [35–37] or a unique exposition in [38, Chapter 9] from a more computer science perspective.

Perhaps the simplest way to explain quantum mechanics is to actually "take the physics out of it" [38] and approach it from the perspective of probability distribution. Consider a discrete probability distribution $p(X)$ for some random variable X that takes values from a finite set $\{x_1, x_2, \cdots, x_n\}$. Then we could write down p as a vector $\mathbf{p} = (p_1, p_2, \cdots, p_n)$ where $p_i = p(X = x_i)$. In our usual "classical" probability theory we require that the probabilities p_i be real numbers from 0 to 1 and $p_1 + p_2 + \cdots + p_n = 1$. In quantum mechanics, the state of a physical system could also be described as a vector $\boldsymbol{\psi} = (\psi_1, \psi_2, \cdots, \psi_n)$, called the *quantum state*, and each element of the vector is a *complex* number that corresponds to a possible state of the system. The probability of a system being in a particular state i is given by $|\psi_i|^2$. Hence unlike

the usual probability theory, our requirement for the state vector is that ψ_i be any *complex* number and $|\psi_1|^2+|\psi_2|^2+\cdots+|\psi_n|^2 = 1$. If we think of the usual probability theory as a theory that requires the 1-norm of the state vector $\mathbf{p}$ to be unity, we could regard quantum mechanics as a new probability theory where the *2-norm* of the state vector normalizes to unity.

A quantum state $\boldsymbol{\psi}$ belongs to the space of unit vectors in $\mathbb{C}^N$, which is called the N-dimensional *Hilbert space*. In quantum physics instead of the usual notations for vectors, we use the *Dirac braket notation* to represent quantum states. We use the symbol

$$|\psi\rangle$$

to describe a quantum state represented by a unit column vector in $\mathbb{C}^N$, while $\langle\psi|$ is its conjugate transpose, or dual vector in the Hilbert space. As a simple example, if we consider the quantum state in our old notation $\boldsymbol{\psi} = \binom{\alpha}{\beta}$ where $\alpha, \beta \in \mathbb{C}$ and $\|\boldsymbol{\psi}\|_2 = |\alpha|^2 + |\beta|^2 = 1$, in our new notation we have

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle, \tag{1.1}$$

where $|0\rangle = \binom{1}{0}$ and $|1\rangle = \binom{0}{1}$. When both α and β are non-zero, the state vector represents a physical system that is in the state $|0\rangle$ and $|1\rangle$ *at the same time*. This uniquely quantum mechanical phenomenon is called *superposition*.

The conjugate transpose of $|\psi\rangle$ defined in (1.1) then becomes $\langle\psi| = \alpha^*\langle 0| + \beta^*\langle 1|$ where $\langle 0| = (1 \quad 0)$ and $\langle 1| = (0 \quad 1)$. This allows us to write the inner product between the state $|\psi\rangle$ and another state $|\phi\rangle = \gamma|0\rangle + \delta|1\rangle$ as $\langle\phi|\psi\rangle = \gamma^*\alpha + \delta^*\beta$. Note that in this braket notation it is the direction of the braket $|\rangle$ or $\langle|$ that matters. Whatever that is inside the bracket is merely a label and serve no operational meaning in calculation. For all purposes one could even have $|\☺\rangle$ and $\langle\☹|$ and they are perfectly valid notations for quantum states. The labelling of $|0\rangle$ and $|1\rangle$ in Equation 1.1 is conventional for describing the state of a quantum bit, or *qubit*, which will be further introduced in Section 1.4.

A quantum state is not only a complex vector in an abstract sense, but also something that produces physically measurable consequences. In quantum mechanics a *measurement* is specified by a particular basis of quantum states. For example if we measure the state in Equation 1.1 in the basis $\{|0\rangle, |1\rangle\}$, we will obtain $|0\rangle$ with probabilty $|\langle 0|\psi\rangle|^2 = |\alpha|^2$ and $|1\rangle$ with probability $|\langle 1|\psi\rangle|^2 = |\beta|^2$. In the context of this dissertation we will only deal with measurement with respect to the basis $|x\rangle$ with x being a binary string. In quantum computing this basis is often called the *computational basis*. After measurement, the quantum state of a system collapses to the state that is obtained. For example if we measure $|\psi\rangle$ in Equation 1.1 and obtain $|0\rangle$, the state of the qubit will no longer be $|\psi\rangle$ but $|0\rangle$ instead. Therefore we could think of measurement as a projection operation that occurs probabilistically according to the square of the norm of elements in the state vector.

In classical physics one specifies the state of a particle at any given time by two variables: position x and momentum p. In quantum mechanics the state of the particle is represented by a vector $|\psi\rangle$ in a Hilbert space. In classical physics, any dynamical variable such as the total energy E (the kinetic energy plus the potential energy) of the particle is a scalar function of x and p. In quantum mechanics, any measurable physical quantity becomes Hermitian operators. Of central importance is the Hermitian operator corresponding to the total energy of the system, called the *Hamiltonian*. Since we have introduced quantum states as vectors in a Hilbert space, the Hamiltonian of a physical system is a Hermitian matrix of the same dimensions as the Hilbert space where the quantum states dwell. The Hamiltonian operator is important because it contains essential information about the physical interactions that the quantum system involves. Also the Hamiltonian governs the time evolution of quantum states according to the *Schrödinger equation*[7]:

$$i\frac{\partial}{\partial t}|\psi(t)\rangle = \mathbf{H}|\psi(t)\rangle \tag{1.2}$$

[7]Typically there is a factor $\hbar$ called Planck's constant that multiplies the left side of the equation (see Equation 1.50). But for our purpose it is fine to assume that this constant is 1.

where t is the time variable and $i = \sqrt{-1}$. If the Hamiltonian is time-independent, we can solve Equation 1.2 by direct integration and yield

$$|\psi(t)\rangle = e^{-i\mathbf{H}t}|\psi(0)\rangle. \tag{1.3}$$

Since $\mathbf{H}$ is Hermitian, the operator $\mathbf{U} = e^{-i\mathbf{H}t}$ is unitary. So here we have a minimalist picture of the quantum world[8]: a collection of particles whose states are described by a unit complex vector in a Hilbert space and their interactions described by a Hermitian matrix called the Hamiltonian. The states of these particles are perpetually evolving under the unitary operation dictated by the Hamiltonian. When measured in a basis where the state of the quantum system does not align with any of the basis states, the state vector of the system collapses to one of the basis states according to a probability that is equal to the norm square of the projection (inner product) between the state vector and the said basis state.

The unitary nature of quantum evolution could also be understood by resorting to our earlier comparison to classical probability theory. Consider a transformation of one discrete probability distribution $\mathbf{p}$ to another one $\mathbf{p}'$. If we assume $\mathbf{p}' = \mathbf{M}\mathbf{p}$ for some matrix $\mathbf{M}$, then $\mathbf{M}$ must be a *stochastic matrix*[9]. For the case of quantum states, the general form of matrices that map a quantum state vector $\boldsymbol{\psi}$ to another one $\boldsymbol{\psi}'$ is the unitary matrices.

From the Schrödinger equation 1.2 we could uncover two aspects of quantum mechanics. One obviously deals with the *dynamics* of the physical system, namely how a quantum state evolves over time. The other deals with the *static* aspect which is completely specified by the spectrum of the Hamiltonian $\mathbf{H}$. If our quantum system

[8]In fact quantum states that can be described using a single unit vector is only a specific type of states called *pure states*. In general a quantum state can be not only a unit vector in a Hilbert space, but a probabilistic mixture of unit vectors. That is, a quantum state could be considered as an ensemble of possible states $|\phi_1\rangle$, $|\phi_2\rangle$, $\cdots$ each of which is assigned a probability p_i. Then we use a *density matrix* $\rho = \sum_i p_i|\phi_i\rangle\langle\phi_i|$ to represent the quantum state (note that $|\phi\rangle\langle\phi|$ is an outer product of a vector with its conjugate transpose). Also, in a strict sense Equation 1.2 describes only a *closed* quantum system which has no energy or mass transfer with its external environment. For open systems the gonverning equation will need to be modified from Equation 1.2. Much is to be said about the fully general description of quantum systems but for the purpose of this dissertation it suffices to focus only on pure states and closed systems.

[9]A matrix is stochastic if and only if its elements are non-negative and each row sums up to 1.

is initialized to an eigenvector (called an *eigenstate*) of $\mathbf{H}$ *i.e.* $|\psi(0)\rangle = |\phi\rangle$ for some $|\phi\rangle$ such that $\mathbf{H}|\phi\rangle = E|\phi\rangle$ where the E is the eigenvalue corresponding to $|\phi\rangle$, solving the Schrödinger equation gives

$$|\psi(t)\rangle = e^{-iEt}|\phi\rangle, \tag{1.4}$$

which is but the eigenstate $|\phi\rangle$ multiplied by a time-dependent phase factor e^{-iEt}. We remark that for any quantum state $|\psi\rangle$, multiplying the state vector by a complex phase factor $e^{i\varphi}$ does not change any measurable properties of the quantum state because the probability of finding the system at any possible state remains the same[10]. Therefore from Equation 1.4 we see that when a physical system starts out in an eigenstate of $\mathbf{H}$, it essentially stays in the same state at any time. This has gained the eigenstates of a Hamiltonian a special status in quantum mechanics as the stationary states of a physical system. In particular, the *ground state* of a quantum system, which is the eigenstate of $\mathbf{H}$ with the smallest energy E (called the *ground state energy*), is of even greater importance. We will elaborate this in Section 1.6 in the context of quantum generalizations of computational complexity.

From our discussions so far we see that once we know the Hamiltonian $\mathbf{H}$ of a physical system, we have all the information needed for computing properties of a quantum system, be it static or dynamic, from Schrödinger equation. However, we are met with the most crucial obstacle that prevents us from exactly solving Schrödinger equation in an overwhelmingly majority of physical systems under consideration. It is an obstacle so formidable that the past century of quantum physics and quantum chemistry is almost entirely dedicated to finding heuristic methods that seem to alleviate its burden but without theoretical guarantee. As it turns out[11], for two quantum systems 1 and 2 each in a state $|\psi_1\rangle$ and $|\psi_2\rangle$ respectively, the joint quantum state

[10]For example, consider $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$ as in Equation 1.1 and $|\psi'\rangle = e^{i\varphi}$. The probability of the system being in the state $|0\rangle$ for both states are $|\alpha|^2$.

[11]This is indeed a great mystery of quantum mechanics. There is no deeper theory that explains the tensor product relation but it seems to stand up to the scrutiny of countless experimental observations.

of the combined system is the *tensor product* of the two states. For example, if we consider the states of two qubits each in a state

$$|\psi_1\rangle = \begin{pmatrix} \alpha_1 \\ \beta_1 \end{pmatrix}, \quad |\psi_2\rangle = \begin{pmatrix} \alpha_2 \\ \beta_2 \end{pmatrix}, \tag{1.5}$$

then the joint state of the two qubits is described by

$$|\psi_{12}\rangle = |\psi_1\rangle \otimes |\psi_2\rangle = \begin{pmatrix} \alpha_1\alpha_2 \\ \alpha_1\beta_2 \\ \beta_1\alpha_2 \\ \beta_1\beta_2 \end{pmatrix}. \tag{1.6}$$

This fact has a dramatic consequence. Imagine we would like to study the quantum behaviour of n qubits. Then the joint state of these qubits will be a vector of dimension 2^n, which is *exponential* as the size of the physical system increases. To put the matter in perspective, a droplet of water contains about 10^{21} molecules. To store the quantum state of even a neglible fraction of it would be impossible even with the most powerful supercomputer in the world.

In accordance with the exponential size of the state vector, the Hamiltonian describing the interactions involved in the n qubits will be a $2^n \times 2^n$ Hermitian matrix. By the sheer size of the matrix, if we try to find the ground state energy by diagonalizing the Hamiltonian, the calculation will easily become infeasible on classical computers if we try to deal with physical systems of even a few dozen qubits. This is precisely the difficulty that Feynman [10] was alluding to in the 1980s and it is only natural to consider physically realizing quantum systems that could simulate the dynamics of quantum systems that are presumably beyond the reach of classical computers. We will discuss this point further in Section 1.4.

1.3 Perturbation theory

The discussion at the end of the last section should provide some evidence on why quantum mechanics is hard. Indeed, so far very few problems in quantum physics

and quantum chemistry can be solved analytically compared with those that seem to be beyond analytical approaches. However, if a problem that cannot be solved analytically looks very similar to one that can, one can find approximate solution by modifying the solution to the solvable problem. Suppose we are asked to find the ground state of a Hamiltonian $\tilde{\mathbf{H}}$ which we do not know how to diagonalize analytically but $\tilde{\mathbf{H}}$ is very "close" to a Hamiltonian $\mathbf{H}$ whose spectrum is known. By "close" we mean that $\tilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ for some perturbation $\mathbf{V}$ whose norm is much smaller than $\mathbf{H}$. Then we can find the approximate spectrum of $\widetilde{\mathbf{H}}$ by starting from that of $\mathbf{H}$ and modify it by considering the influence of $\mathbf{V}$ on the spectrum of $\mathbf{H}$. This is the basic idea underlying *perturbation theory*.

There are several different yet somewhat related formulations of perturbation theory. We will start by introducing the *Rayleigh-Schrödinger* (RS) formalism in Section 1.3.1, which is one of the most commonly used frameworks in quantum physics and quantum chemistry and also the first version of perturbation theory introduced in most quantum mechanics textbooks. In Section 1.3.2 we will introduce *self-energy expansion* which is more commonly used in quantum field theory. This is the central formalism that the results of this dissertation is based on. We also mention its connection to the RS formulation.

1.3.1 Rayleigh-Schrödinger formalism

The first step is to identify the magnitude of perturbation as an expansion parameter. Let $\lambda = \|\mathbf{V}\|_{\mathbf{2}}$ where $\|\cdot\|_2$ is the 2-norm of a matrix. Then $\tilde{\mathbf{H}} = \mathbf{H} + \lambda\hat{\mathbf{V}}$ where $\hat{\mathbf{V}}$ is the perturbation $\mathbf{V}$ normalized with respect to its 2-norm. Our goal to find the eigenpairs of $\tilde{\mathbf{H}}$, namely eigenvalues $\tilde{E}_n$ and eigenstates $|\tilde{n}\rangle$ such that $\tilde{\mathbf{H}}|\tilde{n}\rangle = \tilde{E}_n|\tilde{n}\rangle$. Accordingly we will denote the eigenpairs of $\mathbf{H}$ by $\mathbf{H}|n\rangle = E_n|n\rangle$. We now assume that there is a series expansion for both $\tilde{E}_n$ and $|\tilde{n}\rangle$ in λ:

$$\tilde{E}_n = E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots \tag{1.7}$$

$$|\tilde{n}\rangle = |n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots \tag{1.8}$$

where $E_i^{(j)}$ denotes the j^{th} term in the expansion for the i^{th} eigenvalue (energy level). When $\lambda = 0$, $\tilde{\mathbf{H}} = \mathbf{H}$, we recover the solvable problem. Hence trivially $\tilde{E}_n^{(0)} = E_n$ and $|n^{(0)}\rangle = |n\rangle$. These are our 0^{th} order approximation. To obtain higher order terms in Equations (1.7) and (1.8), we use the relationship $\tilde{\mathbf{H}}|\tilde{n}\rangle = \tilde{E}_n|\tilde{n}\rangle$. Substitute Equations (1.7) and (1.8) together with $\tilde{\mathbf{H}} = \mathbf{H} + \lambda\hat{\mathbf{V}}$ we have

$$\begin{aligned}\left(\mathbf{H} + \lambda\hat{\mathbf{V}}\right)\left(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots\right) \\ = \left(E_n^{(0)} + \lambda E_n^{(1)} + \lambda^2 E_n^{(2)} + \cdots\right)\left(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots\right).\end{aligned} \tag{1.9}$$

Also, from Section 1.2 we mentioned that the state vectors must be unit vectors. Hence $\langle\tilde{n}|\tilde{n}\rangle = 1$. Substituting in the expansion (1.8), we have

$$\left(\langle n^{(0)}| + \lambda\langle n^{(1)}| + \lambda^2\langle n^{(2)}| + \cdots\right)\left(|n^{(0)}\rangle + \lambda|n^{(1)}\rangle + \lambda^2|n^{(2)}\rangle + \cdots\right) = 1. \tag{1.10}$$

Notice that if we take the 0^{th} order term in both brackets, $\langle n^{(0)}|n^{(0)}\rangle = \langle n|n\rangle = 1$. Therefore all terms of non-zero powers of λ on the left side must vanish. Matching powers of λ on both sides of (1.9) and (1.10) we could iteratively compute higher order terms in the perturbative expansion. Take λ^1 for example. From (1.10) we have

$$\langle n^{(0)}|n^{(1)}\rangle + \langle n^{(1)}|n^{(0)}\rangle = 0, \tag{1.11}$$

$$\mathbf{H}|n^{(1)}\rangle + \hat{\mathbf{V}}|n^{(0)}\rangle = E_n^{(1)}|n^{(0)}\rangle + E_n^{(0)}|n^{(1)}\rangle. \tag{1.12}$$

Rearranging terms in (1.12), we have

$$(\mathbf{H} - E_n^{(0)}\mathbf{I})|n^{(1)}\rangle = -\hat{\mathbf{V}}|n^{(0)}\rangle + E_n^{(1)}|n^{(0)}\rangle. \tag{1.13}$$

Here the identity operator $\mathbf{I}$ is often omitted in physics texts but for linear algebraic rigor we keep it. Left multiply both sides of Equation (1.13) by the row vector $\langle n^{(0)}|$, we see that the left side completely vanishes because from (1.11) we have[12] $\langle n^{(0)}|n^{(1)}\rangle = 0$ and $\mathbf{H}$ is diagonal. On the right side of (1.13), multiplying by $\langle n^{(0)}|$ from the left gives $-\langle n^{(0)}|\hat{\mathbf{V}}|n^{(0)}\rangle$ at the first term and $E_n^{(1)}\langle n^{(0)}|n^{(0)}\rangle$ at the second

[12] A subtlety here lies in the fact that multipling by a phase factor $e^{i\varphi}$ on any quantum state $|\psi\rangle$ does not change its physical meaning. Hence the left hand side of (1.11) may not be 0 if we replace $|n^{(1)}\rangle$ by $e^{i\varphi}$. But here we fix φ such that the left hand side is zero, thus giving $\langle n^{(0)}|n^{(1)}\rangle = 0$.

term, which is just $E_n^{(1)}$ since we have $\langle n^{(0)}|n^{(0)}\rangle = 1$. Hence Equation (1.13) gives the first order correction in the n^{th} energy of $\widetilde{\mathbf{H}}$

$$E_n^{(1)} = \langle n^{(0)}|\hat{\mathbf{V}}|n^{(0)}\rangle = \langle n|\hat{\mathbf{V}}|n\rangle. \tag{1.14}$$

To compute the 1^{st} order term in the eigenstate, $|n^{(1)}\rangle$, left multiply the identity operator $\mathbf{I} = \sum_k |k\rangle\langle k|$ on $-\hat{\mathbf{V}}$ on the right hand side of Equation (1.13):

$$\begin{aligned}(\mathbf{H} - E_n^{(0)}\mathbf{I})|n^{(1)}\rangle &= -\textstyle\sum_k |k\rangle\langle k|\hat{\mathbf{V}}|n^{(0)}\rangle + E_n^{(1)} \\ &= -\textstyle\sum_k \langle k|\hat{\mathbf{V}}|n^{(0)}\rangle|k\rangle + \langle n^{(0)}|\hat{\mathbf{V}}|n^{(0)}\rangle|n^{(0)}\rangle \\ &= -\textstyle\sum_{k\neq n}\langle k|\hat{\mathbf{V}}|n^{(0)}\rangle|k\rangle.\end{aligned} \tag{1.15}$$

Multiplying both sides by $\langle m|$, for any $m \neq n$, gives us

$$\begin{aligned}\langle m|(\mathbf{H} - E_n^{(0)}\mathbf{I})|n^{(1)}\rangle &= -\textstyle\sum_{k\neq n}\langle k|\hat{\mathbf{V}}|n^{(0)}\rangle\langle m|k\rangle \\ (E_m - E_n)\langle m|n^{(1)}\rangle &= -\langle m|\hat{\mathbf{V}}|n^{(0)}\rangle \\ \langle m|n^{(1)}\rangle &= -\frac{\langle m|\hat{\mathbf{V}}|n^{(0)}\rangle}{E_m - E_n}\end{aligned} \tag{1.16}$$

where from the first line to the second we have used $\langle m|\mathbf{H} = E_m\langle m|$, $E_n^{(0)} = E_n$ and the orthogonality of the eigenstates $\langle m|k\rangle = \delta_{mk}$. From the second line to the third we have assumed that the subspace with energy (eigenvalue) E_n has no degeneracy, namely $E_m \neq E_n$ for any $m \neq n$. Using $|n^{(0)}\rangle = |n\rangle$ we have the first order term in the expansion for $|\tilde{n}\rangle$ as

$$|n^{(1)}\rangle = \sum_{m\neq n} \frac{\langle m|\hat{\mathbf{V}}|n\rangle}{E_n - E_m}|m\rangle. \tag{1.17}$$

Equation (1.17) can be interpreted as summing over the contribution of all transitions caused by the perturbation $\hat{\mathbf{V}}$ from the current unperturbed state $|n\rangle$ to the other eigenstates $|m\rangle$, scaled by the inverse of the energy difference between the current state and the state $|m\rangle$. In the basis of the eigenstates of $\mathbf{H}$, $\langle m|\hat{\mathbf{V}}|n\rangle$ is the matrix element of $\hat{\mathbf{V}}$ on the m^{th} row and n^{th} column. Since the norm of $\hat{\mathbf{V}}$ is 1, the matrix element should also be bounded from above by a constant. The term $\frac{1}{E_n - E_m}$ reflects the physical intuition that the perturbation should be more likely to cause transitions to the states $|m\rangle$ whose energy levels are close to the current state $|n\rangle$ than those whose energies are further away from E_n.

One could carry out similar computation using Equations (1.9) and (1.10) to find $E^{(2)}$, $|n^{(2)}\rangle$ and higher order terms. Collecting the λ^2 terms in Equation (1.9) gives

$$\begin{aligned} \mathbf{H}|n^{(2)}\rangle + \mathbf{V}|n^{(1)}\rangle &= E_n^{(2)}|n^{(0)}\rangle + E_n^{(1)}|n^{(1)}\rangle + E_n^{(0)}|n^{(2)}\rangle \\ (\mathbf{H} - E_n^{(0)}\mathbf{I})|n^{(2)}\rangle &= -\hat{\mathbf{V}}|n^{(1)}\rangle + E_n^{(2)}|n^{(0)}\rangle + E_n^{(1)}|n^{(1)}\rangle \\ 0 &= -\langle n^{(0)}|\hat{\mathbf{V}}|n^{(1)}\rangle + E_n^{(2)}. \end{aligned} \tag{1.18}$$

Here the second line is obtained from the first line by rearranging the terms and the third line is obtained from the second line by left multiplying both sides by $\langle n^{(0)}|$. Using our earlier result in Equation (1.17) we have the 2^{nd} order correction as

$$E_n^{(2)} = \langle n^{(0)}|\hat{\mathbf{V}} \sum_{m\neq n} \frac{\langle m|\hat{\mathbf{V}}|n\rangle}{E_n - E_m}|m\rangle = \sum_{m\neq n} \langle n|\hat{\mathbf{V}}|m\rangle \cdot \frac{1}{E_n - E_m} \cdot \langle m|\hat{\mathbf{V}}|n\rangle. \tag{1.19}$$

We could interpret the expression for $E_n^{(2)}$ as summing over all processes that starts from $|n\rangle$ and makes a transition to $|m\rangle$ as a first step, contributing the factor $\langle n|\hat{\mathbf{V}}|m\rangle$ with the new energy level E_m contributing a factor $\frac{1}{E_n - E_m}$, and as the second step, returning to $|n\rangle$ while contributing a factor $\langle m|\hat{\mathbf{V}}|n\rangle$. The 2^{nd} order correction to the energy $E_n^{(2)}$ could then be thought of as the total contribution of all processes as such for any $m \neq n$. This type of interpretations intuitively resemble the notions of Motzkin walk that will be used for contructing the algorithms in this dissertation.

Following similar ideas one could compute higher order corrections $E_n^{(3)}$, $E_n^{(4)}$ etc to the series expansion for $\tilde{E}_n$ as well as $|n^{(2)}\rangle$, $|n^{(3)}\rangle$ etc to the series expansion for $|\tilde{n}\rangle$. The detailed computations are rather involved and we will not elaborate on them here. This formalism of perturbation theory is used widely in quantum physics and quantum chemistry. However, a common usage is to compute the first few orders in the expansion and compare the results to experimental measurements (energy and eigenstates are both physically measurable), without much consideration on whether the series converges or if it converges, whether it converges to the exact eigenvalue of $\widetilde{\mathbf{H}}$. In Section 1.2 we have argued that finding the exact eigenvalues of $\tilde{\mathbf{H}}$ is unrealistic in most cases, which is the reason why perturbation theory is needed in the first place. Hence the second best thing one could hope for is to find a converging series of

approximations with known error bound. So that if we do not know the exact value, we could at least have the theoretical guarantee that our answers lie somewhere in an interval that shrinks with every new order of approximation computed. Even this turns out to be too ideal for practical calculations, mostly because of two difficulties:

1. The regime of convergence for perturbation theory is in general hard to pinpoint;
2. If the perturbation expansion is converging, it is hard to obtain an error bound that is tight enough for practical purposes.

Much theoretical progress has been made on the first issue (see [39,40] for example) while the second issue remains on an empirical basis. The error in the perturbative expansion is often obtained by comparing calculation with experimental measurements. Alternatively, one gets a qualitative sense of the error by observing the trend of how the results change as higher order corrections are applied. Since part of the practical motivations for quantum mechanical calculation is to avoid the need to actually perform experiments and the results of trend observing are not necessarily rigorous as theoretical predictions, there is a need for rigorous methods to estimate the error in perturbative expansion. As will be discussed later, this need is even more pressing for adiabatic simulation on a quantum computer and it is the theme of Chapter 4 to address this issue.

1.3.2 Self-energy expansion

The Rayleigh-Schrödinger perturbation theory is useful for finding approximations to the energies and eigenstates of a Hamiltonian $\tilde{\mathbf{H}}$. In this dissertation, however, we are more concerned about approximating the *low-energy effective Hamiltonian* of $\tilde{\mathbf{H}}$, namely $\tilde{\mathbf{H}}_{<E'} = \sum_n{}' \tilde{E}_n |\tilde{n}\rangle\langle\tilde{n}|$ where the summation is over all n such that the energy $\tilde{E}_n$ is less than some predefined cutoff E'. In Section 1.3.1 we have developed perturbative expansions for approximating the energy (eigenvalue) $\tilde{E}_n$ and eigenstate $|\tilde{n}\rangle$. Here we develop a perturbative expansion for the operator $\tilde{\mathbf{H}}_{-}$. We still retain

the assumption that $\tilde{\mathbf{H}} = \mathbf{H}+\mathbf{V}$ where the norm of $\mathbf{V}$ is much smaller than that of $\mathbf{H}$. We divide the Hilbert space $\mathcal{H}$ that $\mathbf{H}$ acts on into two parts, namely $\mathcal{H} = \mathcal{L}_- \oplus \mathcal{L}_+$, where

$$\begin{aligned} \mathcal{L}_- &= \text{span}\{|n\rangle | E_n < E'\} \\ \mathcal{L}_+ &= \text{span}\{|n\rangle | E_n > E'\}. \end{aligned} \tag{1.20}$$

With this partitioning we could write both $\mathbf{H}$ and $\mathbf{V}$ in block forms

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_- & \\ & \mathbf{H}_+ \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} \mathbf{V}_- & \mathbf{V}_{-+} \\ \mathbf{V}_{+-} & \mathbf{V}_+ \end{pmatrix}. \tag{1.21}$$

In order to develop our perturbation expansion it is necessary to define *operator-valued resolvents* $\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$ and $\tilde{\mathbf{G}}(z) = (z\mathbf{I} - \tilde{\mathbf{H}})^{-1}$ where z is a complex parameter. Our perturbative expansion for the low-energy effective Hamiltonian of $\tilde{\mathbf{H}}$ is defined by the *self-energy* operator

$$\mathbf{\Sigma}_-(z) = z\mathbf{I} - [\tilde{\mathbf{G}}_-(z)]^{-1}. \tag{1.22}$$

To derive an expansion from Equation (1.22) we need some linear algebra. Consider first that by definition,

$$\tilde{\mathbf{G}}(z) = \begin{pmatrix} z\mathbf{I}_- - \mathbf{H}_- - \mathbf{V}_- & -\mathbf{V}_{-+} \\ -\mathbf{V}_{+-} & z\mathbf{I}_+ - \mathbf{H}_+ - \mathbf{V}_+ \end{pmatrix}^{-1} = \begin{pmatrix} \tilde{\mathbf{G}}_-(z) & \tilde{\mathbf{G}}_{-+}(z) \\ \tilde{\mathbf{G}}_{+-}(z) & \tilde{\mathbf{G}}_+(z) \end{pmatrix}. \tag{1.23}$$

From the identity

$$\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}^{-1} = \begin{pmatrix} (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & -\mathbf{A}^{-1}\mathbf{B}(\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}(\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1} & (\mathbf{D} - \mathbf{C}\mathbf{A}^{-1}\mathbf{B})^{-1} \end{pmatrix} \tag{1.24}$$

we then have

$$\tilde{\mathbf{G}}_- = (\mathbf{G}_-^{-1} - \mathbf{V}_- - \mathbf{V}_{-+}(\mathbf{G}_+^{-1} - \mathbf{V}_+)^{-1}\mathbf{V}_{+-})^{-1}. \tag{1.25}$$

Using the series expansion $(\mathbf{I}-\mathbf{A})^{-1} = \mathbf{I}+\mathbf{A}+\mathbf{A}^2+\cdots$ on Equation (1.25) and substitute it into Equation (1.22), after some calculation we could obtain the expansion that we have sought:

$$\begin{aligned} \mathbf{\Sigma}_-(z) &= \mathbf{H}_- + \mathbf{V}_- + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-} + \cdots \\ &= \mathbf{H}_- + \mathbf{V}_- + \sum_{r=2}^{\infty} \mathbf{T}_r \end{aligned} \tag{1.26}$$

where $\mathbf{T}_r = \mathbf{V}_{-+}\mathbf{G}_+(\mathbf{V}_+\mathbf{G}_+)^{r-2}\mathbf{V}_{+-}$, $r \geq 2$. Equation (1.26) is the central equation used in this dissertation. We now present some intuitive reason why the self-energy expansion can be considered as an approximation to the low-energy effective Hamiltonian of $\tilde{\mathbf{H}}$. Similar to the Rayleigh-Schrödinger formalism presented in Section 1.3.1, we again use the form $\tilde{\mathbf{H}} = \mathbf{H} + \lambda\hat{\mathbf{V}}$ where $\hat{\mathbf{V}}$ is the perturbation $\mathbf{V}$ normalized by some operator norm. Then Equation (1.26) becomes

$$\mathbf{\Sigma}_-(z) = \mathbf{H}_- + \lambda\hat{\mathbf{V}}_- + \lambda^2\hat{\mathbf{V}}_{-+}\mathbf{G}_+\hat{\mathbf{V}}_{+-} + \cdots . \tag{1.27}$$

Let $\mathcal{L}_-$ be spanned by the ground state (assuming no degeneracy) of $\mathbf{H}$ and $\mathcal{L}_+$ be spanned by all the other eigenstates of $\mathbf{H}$. Then we could consider the expansion (recall the notation $\mathbf{H}|n\rangle = E_n|n\rangle$ for eigenstates of $\mathbf{H}$)

$$\tilde{E}_0 = \langle 0|\mathbf{\Sigma}_-(E_0)|0\rangle = E_0^{(0)} + \lambda E_0^{(1)} + \lambda^2 E_0^{(2)} + \lambda^3 E_0^{(3)} + \cdots . \tag{1.28}$$

If we go through with the computation using Equation (1.26), we find that

$$E_0^{(0)} = \langle 0|\mathbf{H}_-|0\rangle = E_0 \tag{1.29}$$

$$E_0^{(1)} = \langle 0|\mathbf{V}_-|0\rangle = \langle 0|\mathbf{V}|0\rangle \tag{1.30}$$

$$E_0^{(2)} = \langle 0|\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-}|0\rangle = \sum_{m\neq 0}\frac{\langle 0|\mathbf{V}|m\rangle\langle m|\mathbf{V}|0\rangle}{E_0 - E_m} \tag{1.31}$$

$$E_0^{(3)} = \langle 0|\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-}|0\rangle = \sum_{m_1,m_2\neq 0}\frac{\langle 0|\mathbf{V}|m_1\rangle\langle m_1|\mathbf{V}|m_2\rangle\langle m_2|\mathbf{V}|0\rangle}{(E_0 - E_{m_1})(E_0 - E_{m_2})}. \tag{1.32}$$

If one compares the above results with the standard RS perturbation theory shown in Section 1.3.1, one sees that the 0^{th}, 1^{st} (see Equation 1.14 and 1.30), and 2^{nd} order (cf. Equation 1.19 and 1.31) energy corrections are identical. At the 3^{rd} order, RS perturbation theory gives an extra term

$$E_{0,RS}^{(3)} = \sum_{m_1,m_2\neq 0}\frac{\langle 0|\mathbf{V}|m_1\rangle\langle m_1|\mathbf{V}|m_2\rangle\langle m_2|\mathbf{V}|0\rangle}{(E_0 - E_{m_1})(E_0 - E_{m_2})} - \langle 0|\mathbf{V}|0\rangle\sum_{m\neq 0}\frac{\langle 0|\mathbf{V}|m\rangle\langle m|\mathbf{V}|0\rangle}{(E_0 - E_m)^2}, \tag{1.33}$$

which indicates that for perturbation with matrix elements $\langle i|V|j\rangle$ all non-negative, the energy expansion due to self-energy could be an overestimate compared with the RS formalism.

A second angle from which to understand $\mathbf{\Sigma}_-(z)$ as an approximation to the low-energy effective Hamiltonian $\tilde{\mathbf{H}}_{<E'}$ is to consider their formal similarity. Define the projectors $\tilde{\mathbf{\Pi}}_- = \sum_{n:\tilde{E}_n<E'} |\tilde{n}\rangle\langle\tilde{n}|$ and $\mathbf{\Pi}_- = \sum_{n:E_n<E} |n\rangle\langle n|$. Then we could write $\mathbf{\Sigma}_-(z)$ and $\tilde{\mathbf{H}}_{<E'}$ as

$$\begin{aligned} \mathbf{\Sigma}_-(z) &= z\mathbf{I} - [\mathbf{\Pi}_-\tilde{\mathbf{G}}(z)\mathbf{\Pi}_-]^{-1}, \\ \tilde{\mathbf{H}}_{<E'} &= z\mathbf{I} - \tilde{\mathbf{\Pi}}_-[\tilde{\mathbf{G}}(z)]^{-1}\tilde{\mathbf{\Pi}}_-. \end{aligned} \tag{1.34}$$

As is explained by Oliveira and Terhal [8], loosely speaking, if $\mathbf{\Sigma}_-(z)$ is roughly constant in some range of z (defined below in Theorem 1.3.1) then $\mathbf{\Sigma}_-(z)$ is playing the role of $\widetilde{\mathbf{H}}_{<E'}$. This was formalized in [7] and improved in [8] where the following theorem is proven (as in [8] we state the case where $\mathbf{H}$ has $E_0 = 0$ and $E_1 = \Delta$ for some positive number Δ as the spectral gap. We use operator norm $\|\cdot\|$ which is defined as $\|\mathbf{M}\| \equiv \max_{|\psi\rangle\in\mathcal{M}} |\langle\psi|\mathbf{M}|\psi\rangle|$ for an operator $\mathbf{M}$ acting on a Hilbert space $\mathcal{M}$):

Theorem 1.3.1 (Adapted from [7, 8]) *Consider a system $\tilde{\mathbf{H}} = \mathbf{H}+\mathbf{V}$. Let $\|\mathbf{V}\| \leq \Delta/2$ where Δ is the spectral gap of $\mathbf{H}$ and let the low and high spectrum of $\mathbf{H}$ be separated by a cutoff $E' = \Delta/2$. Now let there be an effective Hamiltonian $\mathbf{H}_{\mathit{eff}}$ with a spectrum contained in $[a, b]$. If for some real constant $\epsilon > 0$ and $\forall z \in [a-\epsilon, b+\epsilon]$ with $a < b < \Delta/2-\epsilon$, the self-energy $\mathbf{\Sigma}_-(z)$ has the property that $\|\mathbf{\Sigma}_-(z)-\mathbf{H}_{\mathit{eff}}\| \leq \epsilon$, then each eigenvalue $\tilde{E}_j$ of $\widetilde{\mathbf{H}}_{<E'}$ differs to the j^{th} eigenvalue of $\mathbf{H}_{\mathit{eff}}$ by at most ϵ.*

Theorem 1.3.1 provides a rigorous evidence on how $\mathbf{\Sigma}_-(z)$ approximates the low energy effective Hamiltonian. This is very attractive to us because there is a theoretical guarantee on the upper bound in the error $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|$, which circumvents the difficulties mentioned in Section 1.3.1 on the convergence and error estimation for perturbation expansion. In Section 1.6 we will discuss how Theorem 1.3.1 is used in quantum complexity theory and how it can also be used for adiabatic quantum simulation of many-body systems.

1.4 Quantum computing

In Section 1.2 we have introduced a minimalist picture of the quantum world where the state of a quantum system is described by a state vector $|\psi\rangle$ (of exponential size) in Hilbert space that evolves over time via the unitary operator $\mathbf{U} = e^{-i\mathbf{H}t}$ (See Equation 1.4). For the better part of the past century, experimental quantum physicists have sought to understand the quantum interactions in atoms and molecules by making measurements that allow them to construct Hamiltonians $\mathbf{H}$ that best desribe the quantum behaviours of the physical systems. In this context the Hamiltonian $\mathbf{H}$ is supplied by nature and the goal is to uncover its true form. However, as quantum technologies such as laser and microwave have made rapid advances, it is not only possible to *measure* the quantum interactions in a physical system, but also to *control* the quantum interactions under certain circumstances. In other words, there are well-engineered physical systems where the parameters in the Hamiltonian can be varied experimentally. This provides the hardware basis that allows one to manipulate quantum states of a system, which is what is needed for realizing *quantum computing*. There are many ways of using controllable quantum systems for realizing computations that are potentially beyond classical computers. Here we mention three of them - the gate model (Section 1.4.1), the adiabatic model (Section 1.4.2) and the measurement-based model (Section 1.4.3). The gate model is the most standard model of quantum computation that one would encounter in any quantum computing textbooks [41–43]. The adiabatic model is the model that is most relevant to this dissertation. The three models are shown to be equivalent [44], namely one could simulate another efficiently.

The basic unit of quantum information is a *qubit*, which we have introduced in Section 1.2 as the simplest example of a quantum system. The state of a qubit is decribed by a unit complex vector $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$. The joint state of n qubits are

described by a unit vector in $\mathbb{C}^{2^n}$. A particular subclass of n-qubit states is those that can be expressed as a tensor product of n single-qubit states, namely

$$|\phi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots \otimes |\phi_n\rangle. \tag{1.35}$$

Any n-qubit state that can be expressed in the form of Equation (1.35) is called an *unentangled state*, or *product state*, while any n-qubit state that cannot be expressed in this form is called an *entangled state*. For joint states we often use a compact notation $|\phi_1\phi_2\cdots\phi_n\rangle$ for representing states of the form (1.35). For example we use $|010\rangle$ to represent $|0\rangle \otimes |1\rangle \otimes |0\rangle$. The phenomenon that two quantum systems can be in an entangled state (for example $\frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$) is called *entanglement*. Both entanglement and superposition (Section 1.2) are properties that only quantum mechanical systems exhibit and thus are mechanisms with which quantum computers could perform computation beyond what is possible classically.

A unitary operator acting on the state of a single qubit is a 2×2 matrix $\mathbf{U}$ that can always be expressed as a linear combination of $\mathbf{I}$ and the following *Pauli operators*:

$$\mathbf{X} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \mathbf{Y} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \mathbf{Z} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{1.36}$$

Alternatively we could also write the Pauli operators in terms of outer products of the qubit states $|0\rangle$ and $|1\rangle$: $\mathbf{X} = |0\rangle\langle 1| + |1\rangle\langle 0|$, $\mathbf{Y} = -i|0\rangle\langle 1| + i|1\rangle\langle 0|$ and $\mathbf{Z} = |0\rangle\langle 0| - |1\rangle\langle 1|$. This alternative description supplies the usful intuition that $\mathbf{X}$ "flips" a qubit from 0 to 1 and vice versa, $\mathbf{Y}$ not only flips the qubit but also adds a complex phase factor while $\mathbf{Z}$ does not flip the qubit but adds a phase -1 if the current state is $|1\rangle$. These Pauli operators play an important role in later discussions, especially when it comes to operators that act on an n-qubit system. We say an n-qubit operator $\mathbf{U}$ *operates non-trivially on m qubits* when for a particular subset S of m qubits, let $\mathcal{H}_S$ be the Hilbert space of the qubits in S and $\mathcal{H}_{\bar{S}}$ be the Hilbert space of the other $n-m$ qubits, $\mathbf{U}$ can be written as the tensor product of an identity $\mathbf{I}_{\bar{S}}$ acting on $\mathcal{H}_{\bar{S}}$ and some operator $\mathbf{Q}_S \neq \mathbf{I}_S$ acting on $\mathcal{H}_S$. For example, for $n = 5$ qubits, the operator

$$\mathbf{X} \otimes \mathbf{I} \otimes \mathbf{I} \otimes \mathbf{Y} \otimes \mathbf{Z} \tag{1.37}$$

acts non-trivially on qubits 1, 4 and 5. For convenience, we will introduce a subscript notation to represent a single-qubit operator. For any single-qubit operator $\mathbf{A}$, we use $\mathbf{A}_i$ to represent the n-qubit operator $\mathbf{I} \otimes \cdots \otimes \mathbf{I} \otimes \mathbf{A} \otimes \mathbf{I} \otimes \cdots \otimes \mathbf{I}$ where only the i^{th} operator is $\mathbf{A}$ and the other positions are padded with tensor products of 2×2 identity operators $\mathbf{I}$. The example in Equation 1.37 can be written compactly as $\mathbf{X}_1\mathbf{Y}_4\mathbf{Z}_5$.

1.4.1 Gate model

It is mentioned in the opening part of this section that it has become technologically feasible to physically manipulate quantum system under certain conditions. This means that we could potentially apply unitary transformations *of our own choice* to a quantum state instead of the simple free evolution $\mathbf{U} = e^{-i\mathbf{H}t}$, which sets the stage for realizing quantum computation. On an abstract level, we could think of a *quantum algorithm* as a unitary evolution $\mathbf{U}$. To solve a computational problem, we start from an initial state (without loss of generality assume it is $|00\cdots0\rangle$) and apply $\mathbf{U}$ to it, yielding a final state $|\psi\rangle_{\text{final}} = \mathbf{U}|00\cdots0\rangle$ that encodes the solution to the computational problem. For a quantum algorithm involving n qubits, the unitary $\mathbf{U}$ is of size $2^n \times 2^n$. In order to realize $\mathbf{U}$ physically, one usually decomposes it into a sequence of unitaries $\mathbf{U} = \mathbf{U}_\ell\mathbf{U}_{\ell-1}\cdots\mathbf{U}_2\mathbf{U}_1$ where each $\mathbf{U}_i$ acts non-trivially on at most a few qubits (usually two, since experimentally, controllable two-body interactions are more feasible than other many-body interactions). Each of the $\mathbf{U}_i$ operator is called a *quantum gate*. These gates are the elementary building blocks of a *quantum circuit*. In the gate model of quantum computation, an efficient quantum algorithm on n qubits is a quantum circuit of poly(n) quantum gates.

With regard to the concepts of quantum circuit and quantum gates one could draw analogy from classical computation, where an algorithm is essentially a mapping from an initial bit string (without loss of generality assume it is $00\cdots0$) and some final bit string. This mapping could always be reduced to a sequence of elementary logic

operations such as AND, OR and NOT each of which acts on at most two bits. These logic gates come together to form a circuit that implements the algorithm. Due to its diagrammatic appeal, the circuit model is more intuitive than the more abstract construction of Turing machine, although the two are equivalent to each other. The quantum gate model is also known to be equivalent to the model of quantum Turing machine [6]. Also, like in classical computation, we only need a handful of quantum gates in order to realize arbitrary unitary operation with arbitrarily small error. The set of such gates is called a *universal gate set*. A common choice for a universal gate set is the Hadamard gate H, the phase gate S, the $\pi/8$ gate T and the controlled-NOT gate CNOT:

$$\mathtt{H} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad \mathtt{S} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix}, \quad \mathtt{T} = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi/4} \end{pmatrix}, \quad \mathtt{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{1.38}$$

In general for universal quantum computation one must need at least one two-qubit gate such as CNOT.

The existence of a universal gate set has greatly simplified the physical implementation of a quantum algorithm. However, on the experimental side, progress have been slow in realizing scalable quantum computation in the gate model. In 1995, the first quantum logic gate on 2 qubits was first realized with ion trap [45]. Subsequently in 1999 a 3-qubit [46] and 5-qubit [47] NMR quantum computer was demonstrated. In 2001, the famed Shor's algorithm for factoring was implemented on a 7-qubit NMR quantum computer [48] for factoring the number 15. In 2006, a 12-qubit NMR quantum information processor was benchmarked for the first time [49]. Later on in 2011, a 14-qubit register using ions was demonstrated [50]. Since then, we scarcely know of larger scale experimental realization of quantum systems with potentials for universal gate model quantum computation. Of course, the number of qubits is hardly the only measure of progress in this area. In the past years a great deal of experimental results

have deepened our understanding of the engineering issues that needs to be overcome before a scalable gate model quantum computer can be built.

1.4.2 Adiabatic model

The basic idea of the adiabatic quantum computing (AQC) is introduced by Farhi *et al.* [51]. The first step of the framework is to define a Hamiltonian $\mathbf{H}_P$ whose ground state encodes the solution of the computational problem. Then, we initialize a system in the ground state of some beginning Hamiltonian $\mathbf{H}_B$ that is easy to solve classically, and perform the adiabatic evolution

$$\mathbf{H}(s) = (1-s)\mathbf{H}_B + s\mathbf{H}_P. \tag{1.39}$$

Here $s \in [0,1]$ is a time parameter. In this dissertation we only consider time-dependent function $s(t) = t/T$ for total evolution time T, but in general it could be any general functions that satisfy $s(0) = 0$ and $s(T) = 1$. The adiabatic evolution is governed by the Schrödinger equation for time-dependent Hamiltonian (*cf.* Equation 1.2)

$$i\frac{d}{dt}|\psi(t)\rangle = \mathbf{H}(s(t))|\psi(t)\rangle \tag{1.40}$$

where $|\psi(t)\rangle$ is the state of the system at any time $t \in [0,T]$. Let $\pi_i(s)$ be the i^{th} instantaneous eigenstate of $\mathbf{H}(s)$. In other words, let $\mathbf{H}(s)|\pi_i(s)\rangle = E_i(s)|\pi_i(s)\rangle$ for any s. In particular, let $|\pi_0(s)\rangle$ be the instantaneous ground state of $\mathbf{H}(s)$ at s.

According to the adiabatic theorem [37], for s varying sufficiently slow from 0 to 1, the state of the system $|\psi(t)\rangle$ will remain close to the true ground state $|\pi_0(s(t))\rangle$. At the end of the evolution the system is roughly in the ground state of $\mathbf{H}_P$, which encodes the optimal solution to the problem. If the ground state of $\mathbf{H}_P$ is difficult to find (for instance consider the case for Ising spin glass [23], which is NP-complete), then the adiabatic evolution $\mathbf{H}(s)$ could be used as a heuristic for solving the problem.

An important issue associated with AQC is that the adiabatic evolution needs to be slow enough[13] to avoid exciting the system out of its ground state at any point. In

[13]...thus the name *adiabatic* quantum computing.

order to estimate the minimum runtime T needed for the adiabatic computation, we often use criteria based on the minimum gap $\delta_{\min} = \min_{0\le s\le 1}[E_1(s) - E_0(s)]$ between the ground state energy and the first excited state energy of $\mathbf{H}(s)$, and the rate at which the ground state is excited to the first excited state, which is characterized by $\gamma_{01} = \max_{0\le t\le T} \left|\langle \pi_1(s(t))|\frac{d\mathbf{H}}{dt}|\pi_0(s(t))\rangle\right|$. Specifically, according to the Adiabatic Theorem, if the initial state $|\psi(0)\rangle = |\pi_0(0)\rangle$, then the probability that the final state $|\langle \pi_0(1)|\psi(T)\rangle|^2 \ge 1 - \varepsilon^2$ provided $\frac{\gamma_{01}}{\delta_{\min}} \le \varepsilon$. Since $s(t) = t/T$, by chain rule $\frac{d\mathbf{H}}{dt} = \frac{d\mathbf{H}}{ds}\cdot\frac{1}{T}$. Hence $\gamma_{01} = \frac{1}{T}\cdot\max_{0\le s\le 1}|\langle \pi_1(s)|\frac{d\mathbf{H}}{ds}|\pi_0(s)\rangle| = \tilde{\gamma}_{01}/T$. Therefore in order to guarantee that the success probability is no less than $1-\varepsilon^2$, we must have

$$T \ge \frac{\tilde{\gamma}_{01}}{\delta_{\min}\varepsilon}. \tag{1.41}$$

For $s(t)$ being a linear function and $\mathbf{H}_B$, $\mathbf{H}_P$ being fixed, $\frac{d\mathbf{H}}{ds}$ is a constant. Hence $\tilde{\gamma}_{01}$ is a constant. ε is a predefined threshold and is also considered a constant. Therefore the determining factor for the minimum total time required for the adiabatic evolution to be a valid computation process is the inverse of the minimum gap $\delta_{\min}$. In general $\delta_{\min}$ is hard to find since diagonalizing the Hamiltonian is not an option (both because the Hamiltonian is of exponential size and that the purpose of quantum computing is to circumvent the classical difficulty to diagonalize the Hamiltonian in the first place). However, to show that the adiabatic quantum algorithm is efficient, for an n-qubit Hamiltonian $\mathbf{H}(s)$ we require $T = \text{poly}(n)$, implying that we have to show $\delta_{\min} = \frac{1}{\text{poly}(n)}$. Commonly the inverse polynomial scaling of the minimum gap is established by exploiting some structure of $\mathbf{H}(s)$. Alternatively, for certain cases Quantum Monte Carlo methods are also available.

The largest scale implementation of AQC to date is by D-Wave Systems Inc (hereby called D-Wave for short). In the case of D-Wave, the physical process intended as the adiabatic evolution is more broadly called *quantum annealing* (QA). The concepts of QA and AQC are closely related and almost synonymous, with the subtle difference being that in practice the quantum state of a physical system is always a mixed state (see footnote 8, Section 1.2) instead of a pure state $|\pi_0(0)\rangle$ as is assumed in AQC. Hence QA is used as a more general term that is not specific

to pure states. The quantum processors manufactured by D-Wave are essentially a transverse Ising model (TIM) with tunable local fields and coupling coefficients:

$$\mathbf{H}_{\mathrm{TIM}} = \sum_i \Delta_i \mathbf{X}_i + \sum_i h_i \mathbf{Z}_i + \sum_{i,j} J_{ij} \mathbf{Z}_i \mathbf{Z}_j \tag{1.42}$$

where the parameters Δ_i, h_i and J_{ij} are tunable physically. The qubits are connected in a specific graph geometry that allows for embedding of arbitrary graphs as its minor. The adiabatic evolution starts with a beginning Hamiltonian $\mathbf{H}_B = -h \sum_i \mathbf{X}_i$ and finishes with a problem Hamiltonian

$$\mathbf{H}_P = \sum_i h_i \mathbf{Z}_i + \sum_{i,j} J_{ij} \mathbf{Z}_i \mathbf{Z}_j. \tag{1.43}$$

Equation (1.43) describes a classical Ising model, whose ground state is NP-complete to find in the worst case [52, 53]. Therefore we could encode any combinatorial optimization problem in NP into the parameter assignments $\{h_i, J_{ij}\}$ of $\mathbf{H}_P$ and use the adiabatic evolution under $\mathbf{H}(s) = (1-s)\mathbf{H}_B + s\mathbf{H}_P$ as a method for reaching the ground state of $\mathbf{H}_P$. Here the evolution schedule $s(t)$ is typically nonlinear, taking into account the general feature that the minimum spectral gap always occurs in the middle of the evolution and thus the optimal schedule should be fast in the beginning and the end but slow in the middle.

Although it is tempting to consider the adiabatic evolution as a way of solving NP-complete problems, we note that with a few known exceptions [51], the spectral gap $\delta_{\min}$ typically becomes *exponentially small* as the size of the problem instance grows. Hence to find the ground state of $\mathbf{H}_P$ within an error margin ε in the sense of Equation 1.41, in the worst case one needs at least an exponential amount of evolution time. With that being said, the quantum annealing devices manufactured by D-Wave still holds the potential as a heuristic approach that may be able to solve problems of practical interest more efficiently than other classical heuristics because it makes use of phenomena that are unavailable to classical information processing. The extent to which such potential could materialize is currently a subject of intense study.

Since the initial proposal of QA [54–57], there has been much interest in the search for practical problems where it can be advantageous with respect to classi-

cal algorithms [24–27, 57–82], particularly simulated annealing (SA) [83–85]. Extensive theoretical, numerical and expeirmental efforts have been dedicated to studying the performance of quantum annealing on problems such as Satisfiability [51, 86, 87], Exact Cover [56, 87], Max Independent Set [87], Max Clique [88], integer factorization [89], Graph Isomorphism [90, 91], Ramsey number [92], binary classification [93, 94], unstructured search [95] and search engine ranking [96]. Many of these approaches [51, 56, 86, 88–94] recast the computational problem at hand into a problem of finding the ground state of a classical Ising model in the form of Equation 1.43. The progress so far have deepened our understanding of the regimes and mechanisms of quantum speedup [97]. However, it still remains to find definitive evidence that the D-Wave devices provide algorithmic speedup *of practical use*.

Regardless of whether interesting sets of problems can be found on which QA outperforms classical heuristics, the quantum annealing devices constructed to date represent an important step towards large scale quantum information technology. If the current stage is classical Ising model of the form in Equation 1.43, which is universal for classical computation (NP-complete), a logical next stage would be to consider AQC that simulates a *universal quantum computer*. As is already mentioned, AQC is equivalent to the standard circuit model of quantum computing [44]. Specifically, any quantum circuit (Section 1.4.1) of length L can be simulated by an adiabatic quantum computation of poly(L) time that is governed by a slow varying Hamiltonian of the form in Equation 1.39. The basic idea behind the construction of $\mathbf{H}_B$ and $\mathbf{H}_P$ in [44] starts from the proof of a quantum version of Cook-Levin theorem [31]. In [31, Section 14.4] a Hamiltonian construction is proposed for simulating arbitrary quantum circuit $\mathbf{U} = \mathbf{U}_L\mathbf{U}_{L-1}\cdots\mathbf{U}_2\mathbf{U}_1$ where the unitary operators $\mathbf{U}_i$ are individual quantum gates. Let $|\psi_\ell\rangle$, $\ell = 0, 1, \cdots, L$, be the state at the ℓ^{th} step, namely $|\psi_\ell\rangle = \mathbf{U}_\ell|\psi_{\ell-1}\rangle$ and $|\psi_0\rangle$ is the initial state of the circuit. The proof in [31, Section

14.4] then provides a recipe for constructing a Hamiltonian $\mathbf{H_U}$ correponding to the circuit $\mathbf{U}$ with the property that its ground state is[14]

$$|\Psi\rangle = \frac{1}{\sqrt{L+1}} \sum_{\ell=0}^{L} |\psi_\ell\rangle \otimes |1^\ell 0^{L-\ell}\rangle. \tag{1.44}$$

The state $|\Psi\rangle$ in Equation 1.44 is called the *history state* because it is a superposition of all intermediate states $|\psi_\ell\rangle$ entangled with a separate L-qubit register that distinguishes each intermediate state by using different numbers of 1's in its state. We could think of it as a "clock register" that keeps track of how far along the state $|\psi_\ell\rangle$ is in the sequence of quantum gates. The construction of AQC process in [31] starts from identifying $\mathbf{H_U}$ as the problem Hamiltonian $\mathbf{H}_P$ in the adiabatic evolution under $\mathbf{H}(s) = (1-s)\mathbf{H}_B + s\mathbf{H}_P$. This way as the adiabatic evolution starts from the initial state, which is the ground state of $\mathbf{H}_B$, and proceeds gradually towards the ground state of $\mathbf{H}_P$, by adiabatic theorem the final state will approach the history state $|\Psi\rangle$. To read out the final state $|\psi_L\rangle$, we simply measure the clock register and see if we get $|1^L\rangle$, which happens once in $O(L)$ trials on average (see the discussion in Section 1.2 on measurements). For efficient quantum circuits on n qubits $L = \text{poly}(n)$ and hence we are able to simulate the quantum circuit efficiently with the adiabatic quantum process. Of course, the minimal spectral gap $\delta_{\min}$ during the evolution process is also shown to be $1/\text{poly}(n)$ in [44].

The construction of $\mathbf{H_U}$ according to [31, 44] contains several terms in order to ensure that its ground state is uniquely the history state $|\Psi\rangle$. The basic idea is to add different terms in the Hamiltonian so that any quantum state whose form deviates from $|\Psi\rangle$ will receive an energy penalty *i.e.* $\langle\psi|\mathbf{H_U}|\psi\rangle > \langle\Psi|\mathbf{H_U}|\Psi\rangle$ for any $|\psi\rangle$ that deviates from $|\Psi\rangle$. We will not go through each term of the Hamiltonian in detail but to mention that one of the terms is $\mathbf{H}_{\text{prop}} = \mathbf{H}_1 + \mathbf{H}_2 + \cdots + \mathbf{H}_L$ where if we use the condensed label $|\ell\rangle = |1^\ell 0^{L-\ell}\rangle$ for the clock register, each term $\mathbf{H}_i$ is written as

$$\mathbf{H}_\ell = -\frac{1}{2}\mathbf{U}_\ell \otimes |\ell\rangle\langle\ell-1| - \frac{1}{2}\mathbf{U}_\ell^\dagger \otimes |\ell-1\rangle\langle\ell| + \frac{1}{2}\mathbf{I} \otimes (|\ell\rangle\langle\ell| + |\ell-1\rangle\langle\ell-1|). \tag{1.45}$$

[14]Here the notation $|1^a 0^b\rangle$ represents a state of $a+b$ qubits with the first a qubits in the $|1\rangle$ state and the last b qubits in the $|0\rangle$ state.

The basic intuition for $\mathbf{H}_{\text{prop}}$ is that it represents how a quantum state propagates through the quantum circuit. Take the first term for example: it applies $\mathbf{U}_\ell$ onto the qubits that represent those of the quantum circuit and at the same time updates the clock register by applying $|\ell\rangle\langle\ell-1|$ to move it forward from $|\ell-1\rangle$ to $|\ell\rangle$. The second term has the reverse meaning: if the inverse of a quantum gate is applied (since for unitaries $\mathbf{U}^\dagger\mathbf{U}=\mathbf{I}$) then the clock register is updated backwards. This is also to ensure that $\mathbf{H}_i$ is a Hermitian matrix. The last term essentially says that if nothing is done to the qubits in the circuit (*i.e* if the identity $\mathbf{I}$ is applied) then the state of the clock register also remains the same.

As mentioned in Section 1.4.1, to realize universal quantum computing one must use a universal gate set such as the one described in Equation 1.38. Hence some of the gates $\mathbf{U}_\ell$ could be a two-qubit gate. Note from Equation 1.45 that the $\mathbf{U}_\ell$ terms are also coupled to the corresponding operator on the clock register. For instance the operator $|\ell\rangle\langle\ell-1|$ with $|\ell\rangle = |1^\ell 0^{L-\ell}\rangle$ can be realized by simply applying $|1\rangle\langle 0|$ onto the ℓ^{th} qubit in the clock register. Of course we have to restrict to a subspace where all the clock register qubits other than the ℓ^{th} remain in the same state but this is taken care of by other terms in the construction of $\mathbf{H}_{\mathbf{U}}$. From the discussion after Equation 1.36 we can see that $|1\rangle\langle 0|_{\ell,c} = \frac{1}{2}(\mathbf{X}_{\ell,c} - i\mathbf{Y}_{\ell,c})$ where the subscript ℓ, c denotes the ℓ^{th} clock qubit to distinguish it from the qubits in the quantum circuit being simulated, and $i=\sqrt{-1}$. By the same token, if $\mathbf{U}_\ell$ happens to a `CNOT` gate then from Equation 1.38 we have $\mathbf{U}_\ell = |0\rangle\langle 0|\otimes\mathbf{I} + |1\rangle\langle 1|\otimes\mathbf{X} = \frac{1}{2}(\mathbf{I}+\mathbf{Z}_1) + \frac{1}{2}(\mathbf{I}-\mathbf{Z}_1)\mathbf{X}_2$. Thus the term $\mathbf{U}_\ell \otimes |\ell\rangle\langle\ell-1|$ reads

$$\left(\frac{1}{2}(\mathbf{I}+\mathbf{Z}_1) + \frac{1}{2}(\mathbf{I}-\mathbf{Z}_1)\mathbf{X}_2\right)\otimes\frac{1}{2}(\mathbf{X}_{\ell,c} - i\mathbf{Y}_{\ell,c}) = -\frac{1}{4}\mathbf{Z}_1\mathbf{X}_2\mathbf{X}_{\ell,c} + \frac{1}{4}i\mathbf{Z}_1\mathbf{X}_2\mathbf{Y}_{\ell,c} + \cdots, \tag{1.46}$$

where terms such as $\mathbf{Z}_1\mathbf{X}_2\mathbf{X}_{\ell,c}$ and $\mathbf{Z}_1\mathbf{X}_2\mathbf{Y}_{\ell,c}$ are *three-body* terms that are hard to realize using the current technological capability. For a comparison see Equations 1.42 and 1.43, which contain only simple two-body terms that are far more amenable for physical realization. In [44] the authors also proposed a construction with 6-state

particles[15] interacting in a nearest-neighbor two-body fashion. But 6-dimensional particles are also difficult to experimetally realize so we do not mention it in detail here. Subsequent works [7, 8] have drawn from quantum complexity theory to reduce many-body interactions to two-body. It is one of the central themes of this dissertation to improve these existing methods and also propose new methods for constructing physically realizable two-body Hamiltonians that simulates many-body terms like those that arise in Equation 1.46.

1.4.3 Measurement-based model

The measurement-based model of quantum computation (or alternatively called *one-way quantum computing*) is proposed [98, 99] as an alternative to the more common gate model quantum computing. In the measurement-based model, a computation starts by preparing an entangled quantum state (called *cluster state* [100]) and proceeds by making only single-qubit measurements on a subset of qubits in the state. In a cluster state $|\Phi\rangle$, the qubits are arranged on a graph $G(V, E)$ *e.g.* on a square lattice. Here V is the set of vertices and E is the set of edges. The entire resource for the quantum computation is provided initially in the cluster state whose form is *independent* of the algorithm to be performed. Informally one could think that the cluster state contains all the entanglement needed for the quantum algorithm i.e. the sequence of single-qubit measurements. Furthermore, it is shown that the measurement based model is equivalent to the circuit model [99]. In other words, any quantum circuit can be simulated efficiently by preparing a cluster state, which is simple to do using elementary quantum operations [98], and making measurements on it. Since from the discussion in Section 1.2 we see that measurements are in general projective and thus irreversible, the quantum state at the end of the computation is likely no longer a cluster state, thus the name *one-way quantum computing*.

[15] If a qubit is a 2-state particle with two possible states $|0\rangle$ and $|1\rangle$, a 6-state particle has six possible states. It is called a *qudit* with dimension $d = 6$.

The proposal of one-way quantum computing raises a fascinating possibility that there exist physical systems that are intrinsically universal for quantum computation. One may ask whether there are many-body physical systems whose ground state is naturally a cluster state. Indeed, it is shown [100, 101] that the ground state of a Hamiltonian of the following form is a cluster state $|\Phi\rangle$ with a graph $G(V, E)$:

$$\mathbf{H} = -\sum_{v \in V} \mathbf{H}_v, \quad \mathbf{H}_v = \mathbf{X}_v \prod_{u:(u,v)\in E} \mathbf{Z}_u. \tag{1.47}$$

For example, consider G being a square lattice. Then each term $\mathbf{H}_v$ in the Hamiltonian $\mathbf{H}$ in Equation 1.47 for a vertex not on the edge of the lattice is a five-body term[16]

$$\mathbf{X}_v \mathbf{Z}_{u_1} \mathbf{Z}_{u_2} \mathbf{Z}_{u_3} \mathbf{Z}_{u_4} \tag{1.48}$$

where u_i denotes the four neighbors of the vertex v. Given the importance of the cluster state in measurement-based quantum computing, it behooves us to consider how to realize this many-body interaction using realistic two-body physical systems, which is the central topic of this dissertation.

1.5 Quantum simulation

The idea of *quantum simulation* refers to using quantum computers to study quantum systems that are difficult to model using classical computers. Since Feynman's suggestion that quantum computers would be best suited for simulating quantum mechanics [10], quantum simulation has provided a fruitful ground for developing algorithms where quantum computers may be able to solve computational problems beyond the means of classical computation. Early works [102–107] on quantum simulation have provided algorithms for simulating specific quantum systems. Later developments [108–113] also addressed the problems of computing various properties such as the eigenvalues, dynamics and more.

One of the broadest sets of problems in which quantum simulation could find practical advantages is quantum chemistry. In particular, quantum chemists have been

[16] refer to the discussion after Equation 1.37 for explanation on the notations.

concerned about computing the quantum mechanical properties of molecules such as the energy levels their corresponding eigenstates. These computations appears to be hard for classical computers because the cost of directly solving for the eigenvalues of the molecular Hamiltonian grows exponentially with the problem size. At the end of Section 1.2 we have shown the intuitive reason behind such exponential scaling.

1.5.1 Molecular Hamiltonian

Commonly, quantum chemical calculations of molecular properties treat a molecule as a system of electrons and nuclei. For a molecule of N electrons and M nuclei, the molecular Hamiltonian can be written as

$$\begin{aligned}\mathbf{H} &= \mathbf{K}_e + \mathbf{K}_n + \mathbf{V}_{e-n} + \mathbf{V}_{n-n} + \mathbf{V}_{e-e} \\ &= \left(-\frac{\hbar^2}{2}\sum_{j=1}^{N} m_e^{-1}\nabla_j^2\right) + \left(-\frac{\hbar^2}{2}\sum_{j=1}^{M} m_j^{-1}\nabla_j^2\right) \\ &+ \left(-\sum_{j=1}^{M}\sum_{k=1}^{N}\frac{q_j q_e}{|\mathbf{r}_k - \mathbf{R}_j|}\right) + \left(\sum_{j=1}^{M}\sum_{k=j}^{M}\frac{q_j q_k}{|\mathbf{R}_k - \mathbf{R}_j|}\right) + \left(\sum_{j=1}^{N}\sum_{k=j+1}^{N}\frac{q_e^2}{|\mathbf{r}_j - \mathbf{r}_k|}\right).\end{aligned} \tag{1.49}$$

Before unraveling the terms in Equation 1.49, it is necessary to mention first that the version of quantum mechanics presented in Section 1.2 based on linear algebra is only one formulation of quantum mechanics, largely due to Heisenberg (in fact Heisenberg himself called it "matrix mechanics"). There is another slightly earlier formulation of quantum mechanics by Schrödinger which describes quantum states as continuous *wave functions* $\psi(\mathbf{r})$ with $\mathbf{r} \in \mathbb{R}^3$ being the coordinate of the particle in the space. The wave function is a complex function that satisfies the normalization condition $\int_{\mathbb{R}^3} |\psi(\mathbf{r})|^2 d\mathbf{r} = 1$. Colloquially this formulation was called "wave mechanics". Both formulations are equivalent to each other, since we could introduce an orthonormal[17] set of basis functions $\{\phi_i(\mathbf{r})\}$ and expand any wave function ψ as a linear combination of the basis functions $\psi = \sum_i c_i\phi_i$ where $c_i = \int_{\mathbb{R}^3} \phi_i^*(\mathbf{r})\psi(\mathbf{r})d\mathbf{r}$ is

[17]Here we are working with the space of continuous functions where the inner product between two functions $f, g : \mathbb{R}^3 \mapsto \mathbb{C}$ is defined as $\int_{\mathbb{R}^3} f^*(\mathbf{r})g(\mathbf{r})d\mathbf{r}$.

the inner product between ψ and ϕ_i. This way we have recovered the vector form of the wave function $(c_1, c_2, \cdots)^T$.

Similar to the matrix formulation of quantum mechanics in Section 1.2, in the wave formulation the physical quantities are also represented as operators. The difference is that instead of being matrices, these operators take a continuous form. For example, the momentum operator $\mathbf{P} = -i\hbar\frac{\partial}{\partial x}$ where x is the coordinate in a one-dimensional space. Here $\hbar$ is Planck's constant. Consider the momentum eigenstates $\mathbf{P}\psi(x) = k\psi(x)$. They must satisfy $-i\hbar\frac{\partial}{\partial x}\psi(x) = k\psi(x)$ which gives $\psi(x) \propto e^{i\frac{k}{\hbar}x}$. This essentially describes a plane wave traveling freely in the space. In classical mechanics the kinetic energy of a particle K is related to its momentum p by $\mathbf{K} = \frac{p^2}{2m}$ where m is the mass of a particle. Analogously in quantum mechanics the kinetic energy operator is given by $\frac{\mathbf{P}^2}{2m}$. For three dimensional space the kinetic energy operator becomes $-\frac{\hbar^2}{2m}\nabla^2$. Hence the terms $\mathbf{K}_e$ and $\mathbf{K}_n$ in Equation 1.49 are the kinetic energy operators of the electrons and nuclei, with m_e being the mass of an electron and m_j being the mass of the nucleus of the j^{th} atom.

Both the electrons and nuclei are electrically charged particles and they interact via *Coulomb interactions*. In general, a point with charge q_1 and another with charge q_2 have an attractive force between them if q_1 and q_2 have opposite signs, the magnitude of the attraction force is given by $F(r) = \frac{q_1 q_2}{r^2}$ where r is the distance between the two charges. If both points have charges of the same sign then they repulse each other with force of the same magnitude. This force between objects with static charges is called the Coulumb force. There is a potential energy associated with this force. For instance if two attracting charges get closer the potential energy is released because by the nature of their attracting force they have the tendency to get close. On the other hand if two repulsing charges get closer the potential energy is built up because this goes against the nature of how these two charges interact. More specifically the change in potential energy from distance r_1 to r_2 should be equal to the amount of work done by the Coulomb force $\int_{r_1}^{r_2} F(r)dr$. If we choose $r = \infty$ as the "reference point", we could define a potential energy for every r by computing the negative of

the work done by moving one of the charges from the reference point to r. This gives the Coulumb potential $V(r) = -\int_{\infty}^{r} F(r')dr' = \frac{q_1 q_2}{r}$. In the wave formulation of quantum mechanics the operator corresponding to the Coulumb potential is simply $V(r)$. In Equation 1.49 the terms $\mathbf{V}_{e-n}$, $\mathbf{V}_{n-n}$ and $\mathbf{V}_{e-e}$ account for the electron-electron, electron-nucleus, and nucleus-nucleus coulumb interactions respectively. Here q_e is the charge of an electron and q_j is the charge of the j^{th} nucleus. $\mathbf{r}_k$ is the coordinate of the k^{th} electron and $\mathbf{R}_j$ is the coordinate of the j^{th} nucleus.

To extract important static features of the molecular system, one needs to find the energy (eigenvalues) and eigenstates (eigenfunctions) of the Hamiltonian $\mathbf{H}$ in Equation 1.49. To understand the dynamic quantum behaviours of a molecular system, it remains to solve the time-dependent Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\psi = \mathbf{H}\psi \tag{1.50}$$

where $\mathbf{H}$ is given in Equation 1.49. Here the time-dependent wave function ψ is for all N electrons and M nuclei, namely $\psi = \psi(\{\mathbf{r}_i\}_{i=1}^{N}, \{\mathbf{R}_j\}_{j=1}^{M}, t)$. Equation (1.50) is arguably a coupled differential equation that is very difficult to solve in its general form. One way to proceed is to consider simplifications of Equation 1.49 based on intuitions about the physics. An important physical intuition is that a nucleus is much heavier than an electron. The nucleus mass m_i is typically on the order of 1000 times that of the electron mass m_e. Therefore if we manage to solve the Schrödinger equation (1.50), we are expected to find that the nuclei move much slower than the electrons. So we make an approximation by assuming that the nuclei are not moving at all. This is called the *Born-Oppenheimer approximation*. By introducing this assumption our lives are much easier: if we consider the $\mathbf{R}_j$ parameters fixed, in Equation 1.49 the kinetic energy term for nuclei drops out *i.e.* $\mathbf{K}_n = 0$, the nucleus-nucleus interaction term $\mathbf{V}_{n-n}$ becomes constant (which becomes trivial to treat become it only shifts the spectrum of $\mathbf{H}$ without changing any eigenstates), and finally the electron-nucleus

interaction term $\mathbf{V}_{e-n}$ becomes a sum over terms only for each individual electrons. The new molecular Hamiltonian under Born-Oppenheimer approximation reads

$$\begin{aligned}\mathbf{H} &= \mathbf{K}_e + \mathbf{V}_{e-n} + \mathbf{V}_{e-e} \\ &= \sum_{j=1}^{N}\left(-\frac{\hbar^2}{2m_e}\nabla_j^2 - \sum_{k=1}^{M}\frac{q_e q_k}{|\mathbf{r}_j - \mathbf{R}_k|}\right) + \sum_{j=1}^{N}\sum_{k=j+1}^{N}\frac{q_e^2}{|\mathbf{r}_j - \mathbf{r}_k|}.\end{aligned} \tag{1.51}$$

Equation 1.51 looks much simpler than Equation 1.49. Nonetheless the corresponding Schrödinger equation is still hard to solve. This is because the $\mathbf{V}_{e-e}$ term couples the coordinates $\mathbf{r}_i$ of the electrons and renders the Schrödinger equation unseparable.

In our discussion on the equivalence between the wave formulation and matrix formulation of quantum mechanics we have alluded to the idea that by introducing an orthonormal set of basis functions we could recast a problem in one formulation to the other. Indeed, quantum chemists use specific sets of basis functions, called *orbitals*, for finding the eigenvalues of a molecular Hamiltonian. The solution to either the eigenvalue equation $\mathbf{H}|\Phi\rangle = E|\Phi\rangle$ or the time-dependent Schrödinger equation $i\hbar\frac{\partial}{\partial t}|\psi\rangle = \mathbf{H}|\psi\rangle$ could then be expressed as a function of the orbitals.

Apart from being orthonormal, the set of basis functions we use must obey the same physical constraints as the wave function that we seek. So far we have only considered an electron as a point charge. As a matter of fact, besides its spatial coordinate $\mathbf{r}_i$, an electron also has another degree of freedom which is its *spin*[18]. For an electron the state of its spin is a 2-dimensional unit complex vector $|\sigma\rangle = \alpha|\uparrow\rangle + \beta|\downarrow\rangle$. Here the labels $\uparrow$ and $\downarrow$ represent the distinct "spin up" and "spin down" states. Hence a complete description of the state of a single electron should not only involve its spatial coordinate $\mathbf{r} \in \mathbb{R}^3$ but also its spin coordinate $|\sigma\rangle \in \mathbb{C}^2$. This leads to the introduction of *spin orbitals* for describing a single electron state. Each spin orbital takes the form $\psi(x)\boldsymbol{\sigma}$ with $\boldsymbol{\sigma} \in \{|\uparrow\rangle, |\downarrow\rangle\}$ and the inner product of two spin orbitals is simply $(\int \psi_1^*(\mathbf{r})\psi_2(\mathbf{r})d\mathbf{r}) \cdot \boldsymbol{\sigma}_1^*\boldsymbol{\sigma}_2$. We could now construct a set $\{\phi_j(\mathbf{x})\}_{j=1}^m$ of m orthonormal spin orbitals for describing the state of a single electron, where $\mathbf{x}$ contains both the spatial coordinate $\mathbf{r}$ and the spin coordinate $\boldsymbol{\sigma}$.

[18]The spin angular momentum is an intrinsic property of quantum particles, including electrons and atomic nuclei.

To describe the state of N electrons, which is what we are seeking for the Hamiltonian in Equation 1.51, an initial idea would be to use products of single-electron wavefunctions *i.e.* $\{\phi_{j_1}(\mathbf{x}_1)\phi_{j_2}(\mathbf{x}_2)\cdots\phi_{j_N}(\mathbf{x}_N)|j_1, j_2, \cdots, j_N = 1, 2, \cdots, m\}$ because after all, the Hilbert space $\mathcal{H}$ of N-electron wave functions is a tensor product of single electron Hilbert spaces $\mathcal{H}_1 \otimes \mathcal{H}_2 \otimes \cdots \otimes \mathcal{H}_M$. But there are additional structures that we can introduce to our basis set for N electrons. Because electrons belong to a specific type of quantum particles called *fermions*, its many-body wave function must be antisymmetric with respect to variable exchange[19]:

$$f(\cdots, \mathbf{x}_i, \cdots, \mathbf{x}_j, \cdots) = -f(\cdots, \mathbf{x}_j, \cdots, \mathbf{x}_i, \cdots). \tag{1.52}$$

An immediate consequence of the antisymmetry requirement is that in an N-electron basis function $\phi_{j_1}(\mathbf{x}_1)\phi_{j_2}(\mathbf{x}_2)\cdots\phi_{j_N}(\mathbf{x}_N)$, if there is any pair of indices j_k, j_ℓ that are equal, then such function should vanish. This is because of the antisymmetry requirement in Equation 1.52: $f(\cdots, \mathbf{x}_i, \cdots, \mathbf{x}_i, \cdots) = -f(\cdots, \mathbf{x}_i, \cdots, \mathbf{x}_i, \cdots)$. The requirement that the electrons occupy different spin orbitals in any multi-electron wavefunction is termed *Pauli exclusion principle*. Because of this constraint we are left only with antisymmetric basis functions. A general expression that captures the antisymmetry is given by the *Slater determinant*

$$\frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_{j_1}(\mathbf{x}_1) & \phi_{j_2}(\mathbf{x}_1) & \cdots & \phi_{j_N}(\mathbf{x}_1) \\ \phi_{j_1}(\mathbf{x}_2) & \phi_{j_2}(\mathbf{x}_2) & \cdots & \phi_{j_N}(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{j_1}(\mathbf{x}_N) & \phi_{j_2}(\mathbf{x}_N) & \cdots & \phi_{j_N}(\mathbf{x}_N) \end{vmatrix}. \tag{1.53}$$

Note in Equation 1.53 that if we exchange two electrons by swapping two columns of the determinant, we get a negative sign, effectively capturing Equation 1.52.

[19]Such property comes from certain physical requirements due to relativity, which is rigorously established with the spin-statistics theorem [114–116].

1.5.2 Second quantization

Because of the Pauli exclusion principle, each electron i should occupy a unique spin orbital ϕ_{j_i}. Hence we could represent a Slater determinant with more compact representations by introducing an abstract linear vector space called the *Fock space*, where each determinant is represented by an *occupation number vector* $|\mathbf{n}\rangle$ with $\mathbf{n} = n_1 n_2 \cdots n_m$ being an m-bit string. We use $n_i = 0$ to indicate that ϕ_i is occupied by an electron and $n_i = 0$ to indicate that ϕ_i is not occupied. Note that $|\mathbf{n}\rangle$ itself is not a Slater determinant, but rather a representation of a Slater determinant in the Fock space. In this vector space we use the set of vectors $|\mathbf{n}\rangle$ with $\mathbf{n} \in \{0,1\}^m$ as the orthonormal basis, so that two vectors $|\mathbf{k}\rangle$ and $|\mathbf{k}'\rangle$ have inner product that can be expressed as

$$\langle \mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k},\mathbf{k}'} = \delta_{k_1,k_1'}\delta_{k_2,k_2'}\cdots\delta_{k_m,k_m'}. \tag{1.54}$$

Equation 1.54 is consistent with the inner products between the corresponding Slater determinants. Hence we can consider the Fock space as a 2^m-dimensional vector space spanned by the occupation number vectors. Of particular importance is the state $|00\cdots0\rangle$, which corresponds to the state with no electrons at all. This state is called the *vacuum state* and we denote it as $|\text{vac}\rangle$. We will now show how the other non-vacuum states in the Fock state can be constructed from the vacuum state by introducing two new sets of operators: the creation and annihilation operators.

For any spin orbital index $j = 1, \cdots, m$, the *creation operators* are defined by their actions on the occupation number states:

$$\begin{aligned} a_j^\dagger|n_1\cdots n_{j-1}\mathbf{0}n_{j+1}\cdots n_m\rangle &= (-1)^{n_1+n_2+\cdots+n_{j-1}}|n_1\cdots n_{j-1}\mathbf{1}n_{j+1}\cdots n_m\rangle \\ a_j^\dagger|n_1\cdots n_{j-1}\mathbf{1}n_{j+1}\cdots n_m\rangle &= 0, \end{aligned} \tag{1.55}$$

where the j^{th} element of the occupation number vector is highlighted. The first equation in (1.55) raises the occupation number of the j^{th} spin orbital by 1, thereby "creating" an electron at the spin orbital. The second equation in (1.55) attempts to add another electron at the j^{th} spin orbital, which is already occupied by one electron, and makes the occupation number state vanish. This is consistent with the

Pauli exclusion principle mentioned earlier. Therefore any occupation number vector $|\mathbf{n}\rangle$ can be constructed from the vaccum state by

$$|\mathbf{n}\rangle = (a_1^\dagger)^{n_1}(a_2^\dagger)^{n_2}\cdots(a_m^\dagger)^{n_m}|\text{vac}\rangle. \tag{1.56}$$

From the definition 1.55 we could derive the conjugate transpose of a creation operator, which is the *annihilation operator* a_j:

$$\begin{aligned} a_j|n_1\cdots n_{j-1}\mathbf{1}n_{j+1}\cdots n_m\rangle &= (-1)^{n_1+n_2+\cdots+n_{j-1}}|n_1\cdots n_{j-1}\mathbf{0}n_{j+1}\cdots n_m\rangle \\ a_j|n_1\cdots n_{j-1}\mathbf{0}n_{j+1}\cdots n_m\rangle &= 0. \end{aligned} \tag{1.57}$$

Equation 1.57 is consistent with Equation 1.55 in the sense that for any occupation number states $|\mathbf{m}\rangle$ and $|\mathbf{k}\rangle$, we have the relationship $\langle\mathbf{m}|a_j|\mathbf{k}\rangle^* = \langle\mathbf{k}|a_j^\dagger|\mathbf{m}\rangle$. With further computation it could be checked that the creation and annihilation operators satisfy the anticommutation relations

$$[a_i^\dagger, a_j^\dagger]_+ = 0, \qquad [a_i, a_j]_+ = 0, \qquad [a_i^\dagger, a_j]_+ = \delta_{ij}, \tag{1.58}$$

where $[A, B]_+ = AB + BA$ denotes the anti-commutator between the operators A and B. The anticommutation relations are the fundamental properties that underlie all the other algebraic properties of the *second-quantized formalism of quantum mechanics*, where all operators and states can be constructed from a set of elementary creation and annihilation operators. In Equation 1.56 we have shown how to construct any occupation number state using creation operators. For general states in the Fock space $|\mathbf{c}\rangle = \sum_{\mathbf{k}\in\{0,1\}^m} c_\mathbf{k}|\mathbf{k}\rangle$ we simply take a linear combination of Equation 1.56. We now discuss constructing the operators in the molecular Hamiltonian in Equation 1.51 using the creation and annihilation operators.

The first two terms in the molecular Hamiltonian of Equation 1.51, $\mathbf{K}_e + \mathbf{V}_{e-n}$, consists of only one-electron operators. Hence in the Fock space their actions should only cause one electron to change its spin orbital. In other words, if we let $\mathbf{F}_1 = \mathbf{K}_e + \mathbf{V}_{e-n}$ be the one-electron component of the molecular Hamiltonian, then we could write $\mathbf{F}_1$ in the Fock space as

$$\hat{\mathbf{F}}_1 = \sum_{i,j} f_1(i,j)a_i^\dagger a_j. \tag{1.59}$$

where $f_1(i,j)$ is a scalar coefficient function. Here $a_i^\dagger a_j$ represents the process where one electron hops from the i^{th} spin orbital to the j^{th}. By computing the matrix elements $\langle \mathbf{m}|\hat{\mathbf{F}}_1|\mathbf{k}\rangle$ in the Fock space and comparing with Slater-Condon rules [117], we could determine that[20]

$$f_1(i,j) = \int \phi_i^*(\mathbf{x})\hat{\mathbf{F}}_1\phi_j(\mathbf{x})d\mathbf{x}. \tag{1.60}$$

With the specification in Equation 1.60, the matrix elements of the one-electron operator $\hat{\mathbf{F}}_1$ (Equation 1.59) in the Fock space agree with their counterparts in $\mathbf{F}_1$.

The last term in the molecular Hamiltonian of Equation 1.51, $\mathbf{V}_{e-e}$ acts on two electrons and thus may cause two electrons to change their spin orbitals. We use $\mathbf{V}_{e-e} = \mathbf{F}_2(\mathbf{x}_1, \mathbf{x}_2)$ to denote this two-electron component of the Hamiltonian. Similar to Equation 1.59, we could write $\hat{\mathbf{F}}_2$ as

$$\hat{\mathbf{F}}_2 = \sum_{i,j,k,\ell} f_2(i,j,k,\ell)a_i^\dagger a_j^\dagger a_k a_\ell. \tag{1.61}$$

With a similar approach involving Slater-Condon rules, we can arrive at the identification

$$f_2(i,j,k,\ell) = \int\int \phi_i^*(\mathbf{x}_1)\phi_k^*(\mathbf{x}_2)\mathbf{F}_2(\mathbf{x}_1,\mathbf{x}_2)\phi_j(\mathbf{x}_1)\phi_\ell(\mathbf{x}_2)d\mathbf{x}_1 d\mathbf{x}_2. \tag{1.62}$$

The integrals in Equations 1.60 and 1.62 can be computed efficiently. Hence for a fixed set of spin orbitals we could rewrite the molecular Hamiltonian in Equation 1.51 in second-quantized form

$$\hat{\mathbf{H}} = \sum_{i,j} f_1(i,j)a_i^\dagger a_j + \sum_{i,j,k,\ell} f_2(i,j,k,\ell)a_i^\dagger a_j^\dagger a_k a_\ell. \tag{1.63}$$

Note that although we could compute the f_1, f_2 integrals efficiently, the second-quantized Hamiltonian still acts on a vast vector space that is 2^m dimensional. For N electrons we need the set of spin orbitals to at least be able to hold all electrons, thus $m \geq N$. We have gone through the second-quantization but the exponential scaling of computational effort in diagonalizing $\hat{\mathbf{H}}$ as the number of electrons N increases (Section 1.2) still persists.

[20] See for instance [39, Section 1.4] for details.

1.5.3 Mapping to many-body qubit systems

We have reduced a general problem that is continuous *i.e.* finding the spectrum of the molecular Hamiltonian $\mathbf{H}$ in Equation 1.51, to another problem that is discrete in nature *i.e.* diagonalizing the matrix $\hat{\mathbf{H}}$ in Equation 1.63. In the limit of an infinite, complete basis set, the operators $\mathbf{H}$ and $\hat{\mathbf{H}}$ have the same eigenvalues and their corresponding eigenstates are equivalent representations of each other. The second quantized form $\hat{\mathbf{H}}$ is still hard to diagonalize exactly on a classical computer. But it is particularly convenient for realization on a quantum computer. Observe that the setup of the occupation number vectors in the Fock space provides a natural representation on a quantum computer - let each spin orbital be represented by a qubit, where $|0\rangle$ stands for "unoccupied" and $|1\rangle$ stands for "occupied". As for the creation and annihilation operators, we start by introducing the qubit operators

$$\mathbf{Q}_j^+ = |1\rangle\langle 0|_j = \frac{1}{2}(\mathbf{X}_j - i\mathbf{Y}_j), \qquad \mathbf{Q}_j^- = |0\rangle\langle 1|_j = \frac{1}{2}(\mathbf{X}_j + i\mathbf{Y}_j), \tag{1.64}$$

which seems to mimic the "creation" and "annihilation" behaviours of $a_j^\dagger$ and a_j except that they do not obey the anticommutation relations in Equation 1.58. To ensure that the many-body qubit operators mapped from the creation and annihilation operators do obey the anticommutation relations, one could use the Jordan-Wigner transformation [118]

$$\begin{aligned} a_j^\dagger &\mapsto \mathbf{Q}_j^+ \prod_{k=1}^{j} \mathbf{Z}_k \\ a_j &\mapsto \mathbf{Q}_j^- \prod_{k=1}^{j} \mathbf{Z}_k. \end{aligned} \tag{1.65}$$

By referring to Equation 1.36 one could check that the mapping in Equation 1.65 indeed satisfies the anticommutation relations of the creation and annihilation operators. We note that the Jordan-Wigner transformation is not the only method for mapping creation and annihilation operators to spin operators. One could also use the Bravyi-Kitaev transformation [119,120], which is shown to have advantageous properties for adiabatic quantum many-body simulation for quantum chemistry [121].

A mapping from the second-quantized operators to many-body spin operators completes the last step in recasting an arbitrary molecular Hamiltonian to a many-body qubit Hamiltonian. For example, consider a hydrogen molecule H_2, which contains two electrons and two nuclei that are fixed under Born-Oppenheimer approximation (Section 1.5.1). Each electron has its own spatial coordinate described by an orbital, and a spin coordinate which is 2-dimensional. Hence in a minimal basis we need 4 spin orbitals to represent the two-electron states. In a chosen minimal basis, if we write down the second-quantized Hamiltonian and apply Jordan-Wigner transformation, we obtain a 4-qubit many-body Hamiltonian [119]

$$\begin{aligned}\mathbf{H}_{\mathrm{JW}} &= -0.81261\mathbf{I} + 0.171201\mathbf{Z}_0 + 0.171201\mathbf{Z}_1 - 0.2227965\mathbf{Z}_2 - 0.2227965\mathbf{Z}_3 \\ &+ 0.16862325\mathbf{Z}_1\mathbf{Z}_0 + 0.12054625\mathbf{Z}_2\mathbf{Z}_0 + 0.165868\mathbf{Z}_2\mathbf{Z}_1 + 0.165868\mathbf{Z}_3\mathbf{Z}_0 \\ &+ 0.12054625\mathbf{Z}_3\mathbf{Z}_1 + 0.17434925\mathbf{Z}_3\mathbf{Z}_2 - 0.04532175\mathbf{Y}_3\mathbf{Y}_2\mathbf{X}_1\mathbf{X}_0 \\ &+ 0.04532175\mathbf{X}_3\mathbf{Y}_2\mathbf{Y}_1\mathbf{X}_0 + 0.04532175\mathbf{Y}_3\mathbf{X}_2\mathbf{X}_1\mathbf{Y}_0 - 0.04532175\mathbf{Y}_3\mathbf{Y}_2\mathbf{X}_1\mathbf{X}_0.\end{aligned} \tag{1.66}$$

To experimentally realize such many-body Hamiltonian is far from physically feasible. This necessitates one of the central themes of this dissertation, which is to construct realistic two-body systems that simulate the many-body Hamiltonian desired.

1.6 Quantum Hamiltonian complexity

We have discussed how quantum many-body qubit Hamiltonians can arise in multiple important applications for quantum computing, such as universal AQC (Equation 1.46 in Section 1.4.2), measurement-based quantum computing (Equation 1.47 in Section 1.4.3) and quantum simulation of molecular systems (Equation 1.66 in Section 1.5.3). There are other circumstances where such many-body Hamiltonians arise, for example in the quantum loop models describing topological quantum order require Hamiltonians of four-body interactions [122, 123]. However, it is beyond the scope of this introduction to expand on these subjects.

1.6.1 Local Hamiltonian and QMA

The most general notion of a many-body Hamiltonian which encompasses the many-body systems discussed so far is introduced due to the recently emerging field of *quantum Hamiltonian complexity*. We say that an n-qubit Hamiltonian $\mathbf{H}$ is a *k-local Hamiltonian* if $\mathbf{H} = \sum_{i=1}^{m} \mathbf{H}_i$ where each $\mathbf{H}_i$ acts non-trivially on a distinct subset of at most k qubits[21]. For example, the Hamiltian in Equation 1.46 is 3-local, the Hamiltonian in Equation 1.47 is 5-local and the Hamiltonian in Equation 1.66 is 4-local. This definition of k-local Hamiltonian derives its inspiration from the formulation of Circuit Satisfiability Problem (SAT) in classical computational complexity. In an SAT instance of n Boolean variables (bits) there is a collection of clauses each involving at most k bits. For 3-SAT each clause acts on at most 3 bits and the goal is to find a satisfying assignment or an assignment that violates the minimum number of clauses. We could then consider a k-local Hamiltonian as a sum of "quantum clauses" $\mathbf{H}_i$ each acting on at most k qubits, and the ground state $|\psi\rangle$ must minimize the quadratic form $\langle\psi|\mathbf{H}|\psi\rangle = \langle\psi|\mathbf{H}_1|\psi\rangle + \langle\psi|\mathbf{H}_2|\psi\rangle + \cdots + \langle\psi|\mathbf{H}_m|\psi\rangle$. We could consider the quantity $\langle\psi|\mathbf{H}_m|\psi\rangle$ as a measure of how much a given state $|\psi\rangle$ has "violated" the clause $\mathbf{H}_i$. Then finding the ground state is essentially finding the state assignment $|\psi\rangle$ that minimizes the total violations to all of the quantum clauses. In this sense finding the ground state of a local Hamiltonian is analogous to classical SAT. We formalize this quantum generalization of SAT in the following definition of k-LOCAL HAMILTONIAN problem:

Definition 1.6.1 (k-Local Hamiltonian [31]) *Given a k-local Hamiltonian $\mathbf{H} = \sum_{i=1}^{m} \mathbf{H}_i$ acting on n qubits, and threshold values $a, b \in \mathbb{R}^+$ with $b - a \geq 1/p(n)$ for some polynomial p. Let E_0 be the ground state energy of $\mathbf{H}$. Decide whether $E_0 \leq a$ or $E_0 \geq b$.*

Cook-Levin theorem [28,29] is a classic result in computational complexity theory. The proof of the theorem uses clauses of 3 bits to emulate an arbitrary computational

[21] More general definitions of k-local Hamiltonian assumes the quantum particles are general *qudits* [17]. Here it suffices to consider only qubits.

process of a Turing machine in the sense that the clauses are satisfiable if and only if the Turing machine accepts. In the quantum realm, there is a result that is analogous to the Cook-Levin theorem due to Kitaev [31]. The gist of Kitaev's Hamiltonian construction for the quantum Cook-Levin theorem is already presented in Section 1.4.2 when we are discussing universal AQC. If Cook-Levin theorem establishes 3-SAT as the canonical complete problem for the complexity class NP, then the quantum version establishes Local Hamiltonian as the canonical complete problem for the complexity class QMA, which can be considered as a quantum analogue of NP [15]. We formally define QMA as the following (see [16] for details).

Definition 1.6.2 (Quantum Merlin-Arthur) *A promise problem is in* QMA *if and only if for any string* $\mathbf{x} \in \{0,1\}^n$*, there is a poly*(n) *time uniform family of quantum circuits* $\{\mathbf{U}_n\}$ *that takes as its input* $\mathbf{x}$ *as well as a quantum proof* $|\psi\rangle$*, which is a* $p(n)$*-qubit state with* p *being some polynomial, and* $q(n)$ *ancilla qubits*[22] *in the state* $|0\rangle$ *for some polynomial* q*, such that*

1. *(Completeness) If* $\mathbf{x}$ *is a YES instance, there exists a proof* $|\psi\rangle$ *on* $p(n)$ *qubits such that* $\mathbf{U}_n$ *accepts* $(\mathbf{x}, |\psi\rangle)$ *with probability at least 2/3;*

2. *(Soundness) If* $\mathbf{x}$ *is a NO instance, then for all proofs* $|\psi\rangle$ *on* $p(n)$ *qubits,* $\mathbf{U}_n$ *accepts* $(\mathbf{x}, |\psi\rangle)$ *with probability at most 1/3.*

The quantum analogue of the Cook-Levin theorem [31] has established that 5-Local Hamiltonian is QMA-complete. Subsequently 3-Local Hamiltonian is shown to be QMA-complete independently by Kempe and Regev [124], and Nagaj and Mozes [125]. Then Kempe, Kitaev and Regev [7] showed that 2-Local Hamiltonian is QMA-complete. This completes the complexity characterization of k-Local Hamiltonian for all k, since 1-Local Hamiltonian is in P because we could simply minimize each term $\mathbf{H}_i$ individually. Attention is then shifted towards 2-Local Hamiltonian with additional restricted properties. For example, Oliveira

[22]The term *ancilla qubits* refer to qubits that serve as auxiliary variables in a quantum computing process.

and Terhal [8] showed that 2-Local Hamiltonian restricted to nearest neighbor interactions on a 2D grid remains to be QMA-complete. Biamonte and Love [33] showed that the form of the simplest QMA-complete Hamiltonian can be reduced to physically relevant models such as

$$\mathbf{H} = \sum_i h_i \mathbf{Z}_i + \sum_{i<j} J_{ij} \mathbf{Z}_i \mathbf{Z}_j + \sum_{i<j} K_{ij} \mathbf{X}_i \mathbf{X}_j. \tag{1.67}$$

By comparing the Hamiltonian in Equation 1.67 with the transverse Ising Hamiltonian in Equation 1.42 as realized by D-Wave, we can see that the Hamiltonian in Equation 1.67 only contains additional XX terms (and local X operators but they are far easier to realize with current experimental capabilities). However, find the ground state of Hamiltonians of the form in (1.67) can be QMA-complete in the worst case, while for the transverse Ising Hamiltonian in Equation 1.42 it is unlikely to be QMA-complete [126].

1.6.2 Perturbative gadgets

Although the model described in Equation 1.67 contains only physically accessible terms, programming problems into a universal adiabatic quantum computer [33] or an adiabatic quantum simulator [127, 128] involves several types of k-body interactions (for bounded k). To reduce from k-body interactions to 2-body is accomplished through the application of *perturbative gadgets*. Perturbative gadgets were introduced as theorem-proving tools in the context of quantum complexity theory yet their experimental realization currently offers the only path towards universal adiabatic quantum computation. In terms of experimental constraints, an important parameter in the construction of these gadgets is a large spectral gap introduced into the ancilla space as part of a penalty Hamiltonian. This large spectral gap often requires control precision well beyond current experimental capabilities and must be improved for practical physical realizations.

A perturbative gadget consists of an ancilla system acted on by Hamiltonian $\mathbf{H}$, characterized by the spectral gap Δ between its ground state subspace and excited

state subspace, and a perturbation $\mathbf{V}$ which acts on both the ancilla and the system. $\mathbf{V}$ perturbs the ground state subspace of $\mathbf{H}$ such that the perturbed low-lying spectrum of the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ captures the spectrum of the target Hamiltonian, $\mathbf{H}_{\text{targ}}$, up to error ϵ. The purpose of a gadget is dependent on the form of the target Hamiltonian $\mathbf{H}_{\text{targ}}$. For example, if the target Hamiltonian is k-local with $k \geq 3$ while the gadget Hamiltonian is 2-local, the gadget serves as a tool for reducing locality[23]. Also if the target Hamiltonian involves interactions that are hard to implement experimentally and the gadget Hamiltonian contains only interactions that are physically accessible, the gadget becomes a generator of physically inaccesible terms from accessible ones. Apart from the physical relevance to quantum computation, gadgets have been central to many results in quantum complexity theory [32, 33, 129, 130]. Hamiltonian gadgets were also used to characterize the complexity of density functional theory [20] and are required components in current proposals related to error correction on an adiabatic quantum computer [131] and the adiabatic and ground state quantum simulator [127, 128]. Since these works employ known gadgets which we provide improved constructions of here, our results hence imply a reduction of the resources required in these past works.

The first use of perturbative gadgets [7] relied on a 2-body gadget Hamiltonian to simulate a 3-body Hamiltonian of the form $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \cdot \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$ with three auxiliary spins in the ancilla space. Here $\mathbf{H}_{\text{else}}$ is an arbitrary Hamiltonian that does not operate on the auxiliary spins. Further, $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are unit-norm operators and α is the desired coupling. For such a system, it is shown that it suffices to construct $\mathbf{V}$ with $\|\mathbf{V}\| < \Delta/2$ to guarantee that the perturbative self-energy expansion approximates $\mathbf{H}_{\text{targ}}$ up to error ϵ [7, 8, 129]. Because the gadget Hamiltonian is constructed such that in the perturbative expansion $\mathbf{\Sigma}_{-}(z)$ (with respect to the low energy subspace, see Equation 1.26 in Section 1.3.2), only virtual excitations that flip all 3 ancilla bits would have non-trivial contributions in the 1^{st} through 3^{rd} order terms. In [9] Jordan and Farhi generalized the construction

[23]Here use the notion *locality* to mean the value k for a k-local Hamiltonian.

in [7] to a general k-body to 2-body reduction using a perturbative expansion due to Bloch [132]. They showed that one can approximate the low-energy subspace of a Hamiltonian containing r distinct k-local terms using a 2-local Hamiltonian. Two important gadgets were introduced by Oliveira and Terhal [8] in their proof that 2-LOCAL HAMILTONIAN on a square lattice is QMA-complete. In particular, they introduced an alternative 3- to 2-body gadget which uses only one additional spin for each 3-body term as well as a "subdivision gadget" that reduces a k-body term to a $(\lceil k/2 \rceil + 1)$-body term using only one additional spin [8]. These gadgets, which we improve in this dissertation, find their use as the de facto standard whenever the use of gadgets is necessitated. For instance, the gadgets from [8] were used by Bravyi, DiVincenzo, Loss and Terhal [129] to show that one can combine the use of subdivision and 3- to 2-body gadgets to recursively reduce a k-body Hamiltonian to 2-body, which is useful for simulating quantum many-body Hamiltonians. We note that these gadgets solve a different problem than the type of many-body operator simulations considered previously [133, 134] for gate model quantum computation, where the techniques developed therein are not directly applicable to our situation.

While recent progress in the experimental implementation of adiabatic quantum processors [92, 135–137] suggests the ability to perform sophisticated adiabatic quantum computing experiments, perturbative gadgets require very large values of Δ. This places high demands on experimental control precision by requiring that devices enforce very large couplings between ancilla qubits while still being able to resolve couplings from the original problem – even though those fields may be orders of magnitude smaller than Δ. Accordingly, if perturbative gadgets are to be used, it is necessary to find gadgets which can efficiently approximate their target Hamiltonians with significantly lower values of Δ.

1.7 Summary

In this introductory chapter we have introduced the subject of quantum mechanics (Section 1.2) and quantum computing (Section 1.4). In particular we have introduced perturbation theory (Section 1.3), which provides the mathematical framework for the rest of the dissertation. We have also motivated the subject of reducing many-body to two-body quantum interactions from various different contexts: universal AQC (Section 1.4.2), measurement-based quantum computing (Section 1.4.3), and quantum simulation of molecular system (Section 1.5). We then introduced the central tool for accomplishing the locality reduction - perturbative gadgets (Section 1.6.2) in the context of quantum Hamiltonian complexity (Section 1.6.1).

2. IMPROVED PERTURBATIVE GADGETS

Application of the adiabatic model of quantum computation requires efficient encoding of the solution to computational problems into the lowest eigenstate of a Hamiltonian that supports universal adiabatic quantum computation. Experimental systems are typically limited to restricted forms of 2-body interactions. Therefore, universal adiabatic quantum computation requires a method for approximating quantum many-body Hamiltonians up to arbitrary spectral error using at most 2-body interactions. Perturbative gadgets offer the only current means to address this requirement. Although the applications of perturbative gadgets have steadily grown since their introduction, little progress has been made in overcoming the limitations of the gadgets themselves. In this chapter of experimentally motivated theoretical study, we introduce several gadgets which require significantly more realistic control parameters than similar gadgets in the literature. We employ analytical techniques which result in a reduction of the resource scaling as a function of spectral error for the commonly used subdivision, 3- to 2-body and k-body gadgets. Accordingly, our improvements reduce the resource requirements of all proofs and experimental proposals making use of these common gadgets. Next, we numerically optimize these new gadgets to illustrate the tightness of our analytical bounds. Finally, we introduce a new gadget that simulates a YY interaction term using Hamiltonians containing only $\{X, Z, XX, ZZ\}$ terms. Apart from possible implications in a theoretical context, this work could also be useful for a first experimental implementation of these key building blocks by requiring less control precision without introducing extra ancillary qubits.

2.1 Overview

Continuing from our discussion in Section 1.6.2, previous works in the literature [7, 8, 33, 129, 130] choose Δ to be a polynomial function of ϵ^{-1} which is sufficient for yielding a spectral error $O(\epsilon)$ between the gadget and the target Hamiltonian. Experimental realizations however, will require a recipe for assigning the minimum Δ that guarantees error within specified ϵ, which we consider here. This recipe will need to depend on three parameters: (i) the desired coupling, α; (ii) the magnitude of the non-problematic part of the Hamiltonian, $\|\mathbf{H}_{\text{else}}\|$; and (iii) the specified error tolerance, ϵ. For simulating a target Hamiltonian up to error ϵ, previous constructions [8, 129, 130] use $\Delta = \Theta(\epsilon^{-2})$ for the subdivision gadget and $\Delta = \Theta(\epsilon^{-3})$ for the 3- to 2-body gadget. We will provide analytical results and numerics which indicate that $\Delta = \Theta(\epsilon^{-1})$ is sufficient for the subdivision gadget (Sections 2.2 and 2.3) and $\Delta = \Theta(\epsilon^{-2})$ for the 3- to 2-body gadget (Sections 2.4 and 2.5), showing that the physical resources required to realize the gadgets are less than previously assumed elsewhere in the literature.

In our derivation of the Δ scalings, we use an analytical approach that involves bounding the infinite series in the perturbative expansion. For the 3- to 2-body reduction, in Appendix 2.5 we show that complications arise when there are multiple 3-body terms in the target Hamiltonian that are to be reduced concurrently and bounding the infinite series in the multiple-bit perturbative expansion requires separate treatments of odd and even order terms. Furthermore, in the case where $\Delta = \Theta(\epsilon^{-2})$ is used, additional terms which are dependent on the commutation relationship among the 3-body target terms are added to the gadget in order to compensate for the perturbative error due to cross-gadget contributions (Appendix A).

The next result of this chapter, described in Section 2.6, is a 3- to 2-body gadget construction that uses a 2-body Ising Hamiltonian with a local transverse field. This opens the door to use existing flux-qubit hardware [135] to simulate $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{Z}_i \mathbf{Z}_j \mathbf{Z}_k$ where $\mathbf{H}_{\text{else}}$ is not necessarily diagonal. One drawback of this construction

is that it requires $\Delta = \Theta(\epsilon^{-5})$, rendering it challenging to realize in practice. For cases where the target Hamiltonian is diagonal, there are non-perturbative gadgets [138–140] that can reduce a k-body Hamiltonian to 2-body. In this work, however, we focus on perturbative gadgets.

The final result of this chapter in Section 2.8 is to propose a gadget which is capable of reducing arbitrary real-valued Hamiltonians to a Hamiltonian with only XX and ZZ couplings. In order to accomplish this, we go to fourth-order in perturbation theory to find an XXZZ Hamiltonian which serves as an effective Hamiltonian dominated by YY coupling terms. Because YY terms are especially difficult to realize in some experimental architectures, this result is useful for those wishing to encode arbitrary QMA-hard problems on existing hardware. This gadget in fact now opens the door to solve electronic structure problems on an adiabatic quantum computer.

To achieve both fast readability and completeness in presentation, each section from Section 2.2 to Section 2.8 consists of a **Summary** subsection and an **Analysis** subsection. The former is mainly intended to provide a high-level synopsis of the main results in the corresponding section. Readers could only refer to the **Summary** sections on their own for an introduction to the results of the chapter. The **Analysis** subsections contain detailed derivations of the results in the **Summary**.

2.2 Improved subdivision gadget

Summary. The subdivision gadget is introduced by Oliveira and Terhal [8] in their proof that 2-local Hamiltonian on square lattice is QMA-Complete. Here we show an improved lower bound for the spectral gap Δ needed on the ancilla of the gadget. A subdivision gadget simulates a many-body target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \cdot \mathbf{A} \otimes \mathbf{B}$ ($\mathbf{H}_{\text{else}}$ is a Hamiltonian of arbitrary norm, $\|\mathbf{A}\| = 1$ and $\|\mathbf{B}\| = 1$) by introducing an ancilla spin w and applying onto it a penalty Hamiltonian $\mathbf{H} = \Delta|1\rangle\langle 1|_w$ so that its ground state subspace $\mathcal{L}_- = \text{span}\{|0\rangle_w\}$ and its excited

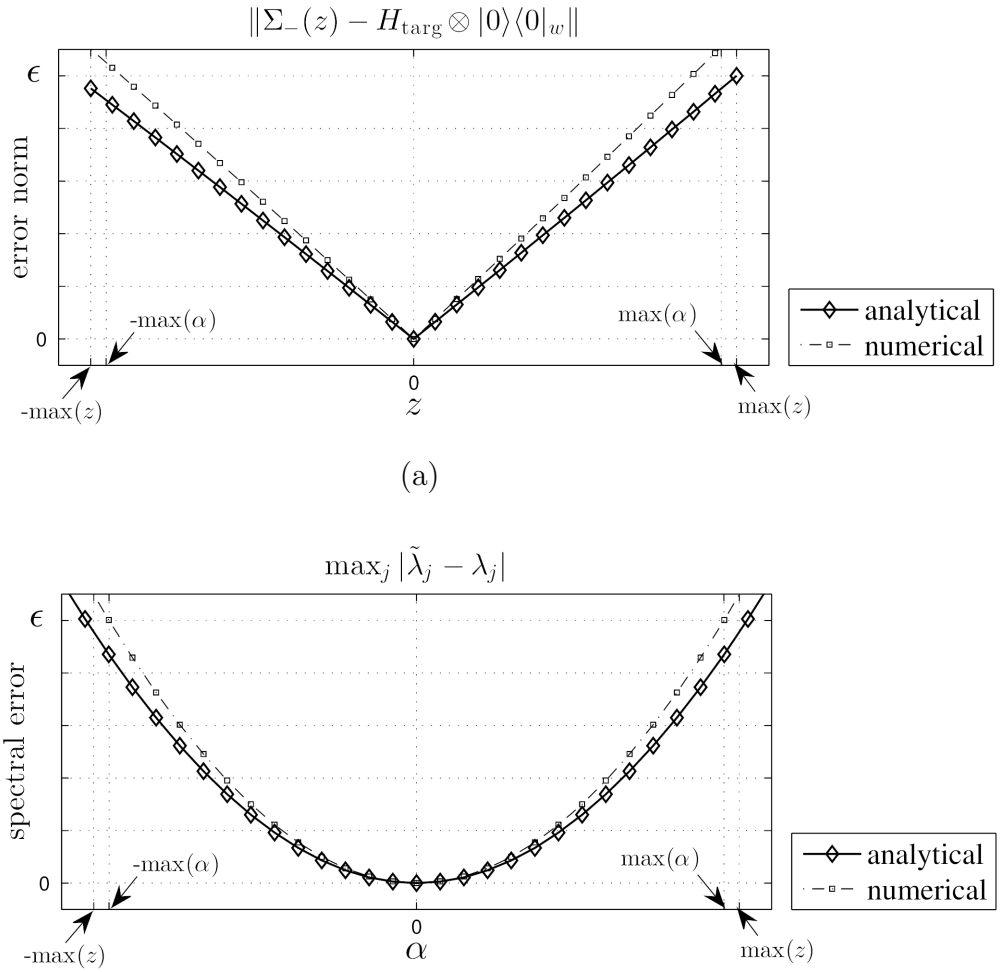


Fig. 2.1. Numerical illustration of the gadget theorem using a subdivision gadget. Here we use a subdivision gadget to approximate $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{Z}_1 \mathbf{Z}_2$ with $\|\mathbf{H}_{\text{else}}\| = 0$ and $\alpha \in [-1, 1]$. $\epsilon = 0.05$. The label "analytical" stands for the case where the value of Δ is calculated using Equation 2.9 when $|\alpha| = 1$. The label "numerical" represents the case where Δ takes the value that yield the spectral error to be ϵ. In (a) we let $\alpha = 1$. $z \in [-\max z, \max z]$ with $\max z = \|\mathbf{H}_{\text{else}}\| + \max \alpha + \epsilon$. The operator $\mathbf{\Sigma}_-(z)$ is computed up to the 3^{rd} order. Subplot (b) shows for every value of α in its range, the maximum difference between the eigenvalues $\tilde{\lambda}_j$ in the low-lying spectrum of $\widetilde{\mathbf{H}}$ and the corresponding eigenvalues λ_j in the spectrum of $\mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$.

subspace $\mathcal{L}_+ = \text{span}\{|1\rangle_w\}$ are separated by energy gap Δ. In addition to the penalty Hamiltonian $\mathbf{H}$, we add a perturbation $\mathbf{V}$ of the form

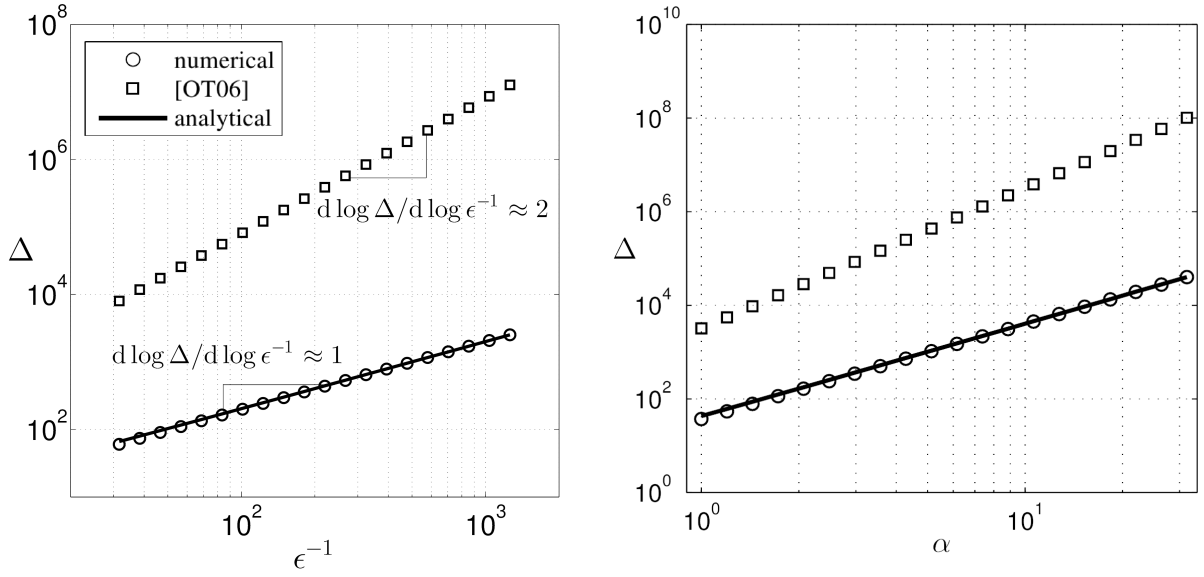
$$\mathbf{V} = \mathbf{H}_{\text{else}} + |\alpha||0\rangle\langle 0|_w + \sqrt{\frac{|\alpha|\Delta}{2}}(\text{sgn}(\alpha)\mathbf{A} - \mathbf{B}) \otimes \mathbf{X}_w. \tag{2.1}$$

Hence if the target term $\mathbf{A} \otimes \mathbf{B}$ is k-local, the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ is at most $(\lceil k/2 \rceil + 1)$-local, accomplishing the locality reduction. Assume $\mathbf{H}_{\text{targ}}$ acts on n qubits. Prior work [8] shows that $\Delta = \Theta(\epsilon^{-2})$ is a sufficient condition for the lowest 2^n levels of the gadget Hamiltonian $\widetilde{\mathbf{H}}$ to be ϵ-close to the corresponding spectrum of $\mathbf{H}_{\text{targ}}$. However, by bounding the infinite series of error terms in the perturbative expansion, we are able to obtain a tighter lower bound for Δ for error ϵ. Hence we arrive at our first result (details will be presented later in this section), that it suffices to let

$$\Delta \geq \left(\frac{2|\alpha|}{\epsilon} + 1\right)(2\|\mathbf{H}_{\text{else}}\| + |\alpha| + \epsilon). \tag{2.2}$$

In Figure 2.2 we show numerics indicating the minimum Δ required as a function of α and ϵ. In Figure 2.2a the numerical results and the analytical lower bound in Equation 2.2 show that for our subdivision gadgets, Δ can scale as favorably as $\Theta(\epsilon^{-1})$. For the subdivision gadget presented in [8], Δ scales as $\Theta(\epsilon^{-2})$. Though much less than the original assignment in [8], the lower bound of Δ in Equation 2.2, still satisfies the condition of Theorem 1.3.1. In Figure 2.2 we numerically find the minimum value of such Δ that yields a spectral error of exactly ϵ.

Analysis. The currently known subdivision gadgets in the literature assume that the gap in the penalty Hamiltonian Δ scales as $\Theta(\epsilon^{-2})$ (see for example [8, 129]). Here we employ a method which uses infinite series to find the upper bound to the norm of the high order terms in the perturbative expansion. We find that in fact $\Delta = \Theta(\epsilon^{-1})$ is sufficient for the error to be within ϵ. A variation of this idea will also be used to reduce the gap Δ needed in the 3- to 2-body gadget (see Section 2.4).



(a) (b)

Fig. 2.2. Comparison between our subdivision gadget with that of Oliveira and Terhal [8]. The data labelled as "numerical" represent the Δ values obtained from the numerical search such that the spectral error between $\mathbf{H}_{\text{targ}}$ and $\widetilde{\mathbf{H}}_{-}$ is ϵ. The data obtained from the calculation using Equation 2.2 are labelled as "analytical". "[OT06]" refers to values of Δ calculated according to the assignment by Oliveira and Terhal [8]. In this example we consider $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{Z}_1 \mathbf{Z}_2$. (a): Gap scaling with respect to ϵ^{-1}. Here $\|\mathbf{H}_{\text{else}}\| = 0$ and $\alpha = 1$. (b): The gap Δ as a function of the desired coupling α. Here $\|\mathbf{H}_{\text{else}}\| = 0$, $\epsilon = 0.05$.

The key aspect of developing the gadget is that given $\mathbf{H} = \Delta|1\rangle\langle 1|_w$, we need to determine a perturbation $\mathbf{V}$ to perturb the low energy subspace

$$\mathcal{L}_- = \text{span}\{|\psi\rangle \otimes |0\rangle_w,\ \ |\psi\rangle \text{ is any state of the system excluding the ancilla spin } w\}$$

such that the low energy subspace of the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ approximates the spectrum of the entire operator $\mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ up to error ϵ. Here we will define V and work backwards to show that it satisfies Theorem 1.3.1. We let

$$\mathbf{V} = \mathbf{H}_{\text{else}} + \frac{1}{\Delta}(\kappa^2\mathbf{A}^2 + \lambda^2\mathbf{B}^2) \otimes |0\rangle\langle 0|_w + (\kappa\mathbf{A} + \lambda\mathbf{B}) \otimes \mathbf{X}_w \tag{2.3}$$

where κ, λ are constants which will be determined such that the dominant contribution to the perturbative expansion which approximates $\widetilde{\mathbf{H}}_{<E_*}$ gives rise to the target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \cdot \mathbf{A} \otimes \mathbf{B}$. In Equation 3.5 and the remainder of the section, by slight abuse of notation, we use $\kappa\mathbf{A} + \lambda\mathbf{B}$ to represent $\kappa(\mathbf{A} \otimes \mathbf{I}_{\mathcal{B}}) + \lambda(\mathbf{I}_{\mathcal{A}} \otimes \mathbf{B})$ for economy. Here $\mathbf{I}_{\mathcal{A}}$ and $\mathbf{I}_{\mathcal{B}}$ are identity operators acting on the subspaces $\mathcal{A}$ and $\mathcal{B}$ respectively. The partitions of $\mathbf{V}$ in the subspaces, as defined in Section 1.3.2 are

$$\begin{gathered}\mathbf{V}_+ = \mathbf{H}_{\text{else}} \otimes |1\rangle\langle 1|_w, \quad \mathbf{V}_- = \left(\mathbf{H}_{\text{else}} + \frac{1}{\Delta}(\kappa^2\mathbf{A}^2 + \lambda^2\mathbf{B}^2)\mathbf{I}\right) \otimes |0\rangle\langle 0|_w, \\ \mathbf{V}_{-+} = (\kappa\mathbf{A} + \lambda\mathbf{B}) \otimes |0\rangle\langle 1|_w, \quad \mathbf{V}_{+-} = (\kappa\mathbf{A} + \lambda\mathbf{B}) \otimes |1\rangle\langle 0|_w.\end{gathered} \tag{2.4}$$

We would like to approximate the target Hamiltonian $\mathbf{H}_{\text{targ}}$ and so expand the self-energy in Equation 1.26 up to 2^{nd} order. Note that $\mathbf{H}_- = 0$ and $\mathbf{G}_+(z) = (z - \Delta)^{-1}|1\rangle\langle 1|_w$. Therefore the self energy $\mathbf{\Sigma}_-(z)$ can be expanded as

$$\begin{aligned}\mathbf{\Sigma}_-(z) &= \mathbf{V}_- + \frac{1}{z-\Delta}\mathbf{V}_{-+}\mathbf{V}_{+-} + \sum_{k=1}^{\infty} \frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}} \\ &= \underbrace{\left(\mathbf{H}_{\text{else}} - \frac{2\kappa\lambda}{\Delta}\mathbf{A} \otimes \mathbf{B}\right) \otimes |0\rangle\langle 0|_w}_{\mathbf{H}_{\text{eff}}} \\ &+ \underbrace{\frac{z}{\Delta(z-\Delta)}(\kappa\mathbf{A} + \lambda\mathbf{B})^2 \otimes |0\rangle\langle 0|_w + \sum_{k=1}^{\infty} \frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}}}_{\text{error term}}.\end{aligned} \tag{2.5}$$

By selecting $\kappa = \text{sgn}(\alpha)(|\alpha|\Delta/2)^{1/2}$ and $\lambda = -(|\alpha|\Delta/2)^{1/2}$, the leading order term in $\mathbf{\Sigma}_-(z)$ becomes $\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$. We must now show that the condition of

Theorem 1.3.1 is satisfied i.e. for a small real number $\epsilon > 0$, $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\| \leq \epsilon, \forall z \in [\min z, \max z]$ where $\max z = \|\mathbf{H}_{\text{else}}\| + |\alpha| + \epsilon = -\min z$. Essentially this amounts to choosing a value of Δ to cause the error term in Equation 2.5 to be $\leq \epsilon$. In order to derive a tighter lower bound for Δ, we bound the norm of the error term in Equation 2.5 by letting $z \mapsto \max z$ and from the triangle inequality for operator norms:

$$\begin{aligned}
\left\| \frac{z}{\Delta(z-\Delta)} (\kappa\mathbf{A} + \lambda\mathbf{B})^2 \otimes |0\rangle\langle 0|_w \right\| &\leq \frac{\max z}{\Delta(\Delta - \max z)} \cdot 4\kappa^2 = \frac{2|\alpha| \max z}{\Delta - \max z} \\
\left\| \sum_{k=1}^{\infty} \frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}} \right\| &\leq \sum_{k=1}^{\infty} \frac{\|\mathbf{V}_{-+}\| \cdot \|\mathbf{V}_+\|^k \cdot \|\mathbf{V}_{+-}\|}{(\Delta - \max z)^{k+1}} \\
&\leq \sum_{k=1}^{\infty} \frac{2|\kappa| \cdot \|\mathbf{H}_{\text{else}}\|^k \cdot 2|\kappa|}{(\Delta - \max z)^{k+1}} \\
&= \sum_{k=1}^{\infty} \frac{2|\alpha|\Delta\|\mathbf{H}_{\text{else}}\|^k}{(\Delta - \max z)^{k+1}}.
\end{aligned} \tag{2.6}$$

Using $\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$, from (2.5) we see that

$$\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \frac{2|\alpha| \max z}{\Delta - \max z} + \sum_{k=1}^{\infty} \frac{2|\alpha|\Delta\|\mathbf{H}_{\text{else}}\|^k}{(\Delta - \max z)^{k+1}} \tag{2.7}$$

$$= \frac{2|\alpha| \max z}{\Delta - \max z} + \frac{2|\alpha|\Delta}{\Delta - \max z} \cdot \frac{\|\mathbf{H}_{\text{else}}\|}{\Delta - \max z - \|\mathbf{H}_{\text{else}}\|}. \tag{2.8}$$

Here going from Equation 2.7 to Equation 2.8 we have assumed the convergence of the infinite series in Equation 2.7, which adds the reasonable constraint that $\Delta > |\alpha| + \epsilon + 2\|\mathbf{H}_{\text{else}}\|$. To ensure that $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \epsilon$ it is sufficient to let expression Equation 2.8 be $\leq \epsilon$, which implies that

$$\Delta \geq \left(\frac{2|\alpha|}{\epsilon} + 1 \right) (|\alpha| + \epsilon + 2\|\mathbf{H}_{\text{else}}\|) \tag{2.9}$$

which is $\Theta(\epsilon^{-1})$, a tighter bound than $\Theta(\epsilon^{-2})$ in the literature [7, 8, 129]. This bound is illustrated with a numerical example (Figure 2.1). From the data labelled as "analytical" in Figure 2.1a we see that the error norm $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|$ is within ϵ for all z considered in the range, which satisfies the condition of the theorem for the chosen example. In Figure 2.1b, the data labelled "analytical" show that the spectral difference between $\widetilde{\mathbf{H}}_{<E_*}$ and $\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ is indeed within ϵ as the theorem promises. Furthermore, note that the condition of Theorem 1.3.1 is only sufficient,

which justifies why in Figure 2.1b for α values at $\max \alpha$ and $\min \alpha$ the spectral error is strictly below ϵ. This indicates that an even smaller Δ, although below the bound we found in Equation 2.9 to satisfy the theorem, could still yield the spectral error within ϵ for all α values in the range. The smallest value Δ can take would be one such that the spectral error is exactly ϵ when α is at its extrema. We numerically find this Δ (up to numerical error which is less than $10^{-5}\epsilon$) and as demonstrated in Figure 2.1b, the data labelled "numerical" shows that the spectral error is indeed ϵ at $\max(\alpha)$ and $\min(\alpha)$, yet in Figure 2.1a the data labelled "numerical" shows that for some z in the range the condition of the Theorem 1.3.1, $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \epsilon$, no longer holds. In Figure 2.1 we assume that ϵ is kept constant. In Figure 2.2a we compute both analytical and numerical Δ values for different values of ϵ.

Comparison with Oliveira and Terhal [8]. We also compare our Δ assignment with the subdivision gadget by Oliveira and Terhal [8], where given a target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \mathbf{Q} \otimes \mathbf{R}$ it is assumed that $\mathbf{Q}$ and $\mathbf{R}$ are operators with finite norm operating on two separate spaces $\mathcal{A}$ and $\mathcal{B}$.

The construction of the subdivision gadget in [8] is the same as the construction presented earlier: introduce an ancillary qubit w with energy gap Δ, then the unperturbed Hamiltonian is $\mathbf{H} = \Delta|1\rangle\langle 1|_w$. In [8] they add a perturbation $\mathbf{V}$ that takes the form of [8, Equation 15]

$$\mathbf{V} = \mathbf{H}'_{\text{else}} + \sqrt{\frac{\Delta}{2}}(-\mathbf{Q} + \mathbf{R}) \otimes \mathbf{X}_w \tag{2.10}$$

where $\mathbf{H}'_{\text{else}} = \mathbf{H}_{\text{else}} + \mathbf{Q}^2/2 + \mathbf{R}^2/2$. Comparing the form of Equation 2.10 and Equation 3.5 we can see that if we redefine $\mathbf{Q} = \sqrt{|\alpha|}\mathbf{A}$ and $\mathbf{R} = \sqrt{|\alpha|}\mathbf{B}$, the gadget formulation is identical to our subdivision gadget approximating $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{A} \otimes \mathbf{B}$ with $\alpha > 0$. In the original work Δ is chosen as [8, Equation 20]

$$\Delta = \frac{(\|\mathbf{H}'_{\text{else}}\| + C_2 r)^6}{\epsilon^2}$$

where $C_2 \geq \sqrt{2}$ and $r = \max\{\|\mathbf{Q}\|, \|\mathbf{R}\|\}$. In the context of our subdivision gadget, this choice of Δ translates to a lower bound

$$\Delta \geq \frac{(\|\mathbf{H}_{\text{else}} + |\alpha|\mathbf{I}\| + \sqrt{2|\alpha|})^6}{\epsilon^2}. \tag{2.11}$$

In Figure 2.2a we compare the lower bound in Equation 2.11 with our lower bound in Equation 2.9 and the numerically optimized Δ described earlier.

2.3 Parallel subdivision and k- to 3-body reduction

Summary. Applying subdivision gadgets iteratively one can reduce a k-body Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \bigotimes_{i=1}^{k} \mathbf{S}_i$ to 3-body. Here each $\mathbf{S}_i$ is a single spin Pauli operator. Initially, the term $\bigotimes_{i=1}^{k} \mathbf{S}_i$ can be broken down into $\mathbf{A} \otimes \mathbf{B}$ where $\mathbf{A} = \bigotimes_{i=1}^{r} \mathbf{S}_i$ and $\mathbf{B} = \bigotimes_{i=r+1}^{k} \mathbf{S}_i$. Let $r = k/2$ for even k and $r = (k+1)/2$ for odd k. The gadget Hamiltonian will be $(\lceil k/2 \rceil + 1)$-body, which can be further reduced to a $(\lceil \lceil k/2 \rceil + 1 \rceil/2 + 1)$-body Hamiltonian in the same fashion. Iteratively applying this procedure, we can reduce a k-body Hamiltonian to 3-body, with the i^{th} iteration introducing the same number of ancilla qubits as that of the many-body term to be subdivided. Applying the previous analysis on the improved subdivision gadget construction, we find that $\Delta_i = \Theta(\epsilon^{-1}\Delta_{i-1}^{3/2})$ is sufficient such that during each iteration the spectral difference between $\widetilde{\mathbf{H}}_i$ and $\widetilde{\mathbf{H}}_{i-1}$ is within ϵ. From the recurrence relation $\Delta_i = \Theta(\epsilon^{-1}\Delta_{i-1}^{3/2})$, we are then able to show a quadratic improvement over previous k-body constructions [129].

Analysis. The concept of parallel application of gadgets has been introduced in [7,8]. The idea of using subdivision gadgets for iteratively reducing a k-body Hamiltonian to 3-body has been mentioned in [8, 129]. Here we elaborate the idea by a detailed analytical and numerical study. We provide explicit expressions of all parallel subdivision gadget parameters which guarantees that during each reduction the error between the target Hamiltonian and the low-lying sector of the gadget Hamiltonian

is within ϵ. For the purpose of presentation, let us define the notions of "parallel" and "series" gadgets in the following remarks.

Remark 2.3.1 (Parallel gadgets) *Parallel application of gadgets refers to using gadgets on multiple terms $\mathbf{H}_{targ,i}$ in the target Hamiltonian $\mathbf{H}_{targ} = \mathbf{H}_{else} + \sum_{i=1}^{m} \mathbf{H}_{targ,i}$ concurrently. Here one will introduce m ancilla spins $w_1, \cdots, w_m$ and the parallel gadget Hamiltonian takes the form of $\widetilde{\mathbf{H}} = \sum_{i=1}^{m} \mathbf{H}_i + \mathbf{V}$ where $\mathbf{H}_i = \Delta|1\rangle\langle 1|_{w_i}$ and $\mathbf{V} = \mathbf{H}_{else} + \sum_{i=1}^{m} \mathbf{V}_i$. $\mathbf{V}_i$ is the perturbation term of the gadget applied to $\mathbf{H}_{targ,i}$.*

Remark 2.3.2 (Serial gadgets) *Serial application of gadgets refers to using gadgets sequentially. Suppose the target Hamiltonian $\mathbf{H}_{targ}$ is approximated by a gadget Hamiltonian $\widetilde{\mathbf{H}}^{(1)}$ such that $\widetilde{\mathbf{H}}_{-}^{(1)}$ approximates the spectrum of $\mathbf{H}_{targ}$ up to error ϵ. If one further applies onto $\widetilde{\mathbf{H}}^{(1)}$ another gadget and obtains a new Hamiltonian $\widetilde{\mathbf{H}}^{(2)}$ whose low-lying spectrum captures the spectrum of $\widetilde{\mathbf{H}}^{(1)}$, we say that the two gadgets are applied in series to reduce $\mathbf{H}_{targ}$ to $\widetilde{\mathbf{H}}^{(2)}$.*

Based on Remark 2.3.1, a parallel subdivision gadget deals with the case where $\mathbf{H}_{\text{targ,i}} = \alpha_i \mathbf{A}_i \otimes \mathbf{B}_i$. α_i is a constant and $\mathbf{A}_i$, $\mathbf{B}_i$ are unit norm Hermitian operators that act on separate spaces $\mathcal{A}_i$ and $\mathcal{B}_i$. Note that with $\mathbf{H}_i = \Delta|1\rangle\langle 1|_{w_i}$ for every $i \in \{1, 2, \cdots, m\}$ we have the total penalty Hamiltonian $\mathbf{H} = \sum_{i=1}^{m} \mathbf{H}_i = \sum_{x\in\{0,1\}^m} h(x)\Delta|x\rangle\langle x|$ where $h(x)$ is the Hamming weight of the m-bit string x. This penalty Hamiltonian ensures that the ground state subspace is $\mathcal{L}_{-} = \text{span}\{|0\rangle^{\otimes m}\}$ while all the states in the subspace $\mathcal{L}_{+} = \text{span}\{|x\rangle | x \in \{0,1\}^m, x \neq 00\cdots 0\}$ receives an energy penalty of at least Δ. The operator-valued resolvent G for the penalty Hamiltonian is (by definition in Section 1.3)

$$\mathbf{G}(z) = \sum_{x\in\{0,1\}^m} \frac{1}{z - h(x)\Delta} |x\rangle\langle x|. \tag{2.12}$$

The perturbation Hamiltonian V is defined as

$$\mathbf{V} = \mathbf{H}_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2 \mathbf{A}_i^2 + \lambda_i^2 \mathbf{B}_i^2) + \sum_{i=1}^{m}(\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i) \otimes \mathbf{X}_{u_i} \tag{2.13}$$

where the coefficients κ_i and λ_i are defined as $\kappa_i = \mathrm{sgn}(\alpha_i)\sqrt{|\alpha_i|\Delta/2}, \lambda_i = -\sqrt{|\alpha_i|\Delta/2}$. Define $\mathbf{P}_- = |0\rangle^{\otimes m}\langle 0|^{\otimes m}$ and $\mathbf{P}_+ = \mathbf{I} - \mathbf{P}_-$. Then if $\mathbf{H}_{\mathrm{targ}}$ acts on the Hilbert space $\mathcal{M}$, $\mathbf{\Pi}_- = \mathbf{I}_{\mathcal{M}} \otimes \mathbf{P}_-$ and $\mathbf{\Pi}_+ = \mathbf{I}_{\mathcal{M}} \otimes \mathbf{P}_+$. Comparing Equation 2.13 with Equation 3.5 we see that the projector to the low-lying subspace $|0\rangle\langle 0|_w$ in Equation 3.5 is replaced by an identity $\mathbf{I}$ in Equation 2.13. This is because in the case of m parallel gadgets $\mathbf{P}_-$ cannot be realized with only 2-body terms when $m \geq 3$.

The partition of V in the subspaces are

$$\begin{aligned}
\mathbf{V}_- &= \left(\mathbf{H}_{\mathrm{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2\mathbf{A}_i^2 + \lambda_i^2\mathbf{B}_i^2)\right) \otimes \mathbf{P}_-, \\
\mathbf{V}_+ &= \left(\mathbf{H}_{\mathrm{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2\mathbf{A}_i^2 + \lambda_i^2\mathbf{B}_i^2)\right) \otimes \mathbf{P}_+, \\
\mathbf{V}_{-+} &= \sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i) \otimes \mathbf{P}_-\mathbf{X}_{u_i}\mathbf{P}_+, \\
\mathbf{V}_{+-} &= \sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i) \otimes \mathbf{P}_+\mathbf{X}_{u_i}\mathbf{P}_-.
\end{aligned} \tag{2.14}$$

The self-energy expansion in Equation 1.26 then becomes

$$\begin{aligned}
\mathbf{\Sigma}_-(z) &= \left(\mathbf{H}_{\mathrm{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2\mathbf{A}_i^2 + \lambda_i^2\mathbf{B}_i^2)\right) \otimes \mathbf{P}_- + \frac{1}{z-\Delta}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)^2 \otimes \mathbf{P}_- \\
&+ \sum_{k=1}^{\infty}\mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^k\mathbf{G}_+\mathbf{V}_{+-}.
\end{aligned} \tag{2.15}$$

Rearranging the terms we have

$$\begin{aligned}
\mathbf{\Sigma}_-(z) &= \underbrace{\left(\mathbf{H}_{\mathrm{else}} + \sum_{i=1}^{m}\left(-\frac{2\kappa_i\lambda_i}{\Delta}\mathbf{A}_i \otimes \mathbf{B}_i\right)\right) \otimes \mathbf{P}_-}_{\mathbf{H}_{\mathrm{eff}}} \\
&+ \underbrace{\left(\frac{1}{\Delta} + \frac{1}{z-\Delta}\right)\sum_{i=1}^{m}(\kappa_i^2\mathbf{A}_i^2 + \lambda_i^2\mathbf{B}_i^2) \otimes \mathbf{P}_-}_{\mathbf{E}_1} \\
&+ \underbrace{\left(\frac{1}{\Delta} + \frac{1}{z-\Delta}\right)\sum_{i=1}^{m}2\kappa_i\lambda_i\mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{P}_-}_{\mathbf{E}_2} + \underbrace{\sum_{k=1}^{\infty}\mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^k\mathbf{G}_+\mathbf{V}_{+-}}_{\mathbf{E}_3}
\end{aligned} \tag{2.16}$$

where the term $\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes \mathbf{P}_-$ is the effective Hamiltonian that we would like to obtain from the perturbative expansion and $\mathbf{E}_1$, $\mathbf{E}_2$, and $\mathbf{E}_3$ are error terms. Theorem 1.3.1 states that for $z \in [-\max(z), \max(z)]$, if $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{targ}} \otimes \mathbf{P}_-\| \leq \epsilon$ then $\widetilde{\mathbf{H}}_{<E_*}$ approximates the spectrum of $\mathbf{H}_{\text{targ}} \otimes \mathbf{P}_-$ by error at most ϵ. Similar to the triangle inequality derivation shown in (2.6), to derive a lower bound for Δ, let $z \mapsto \max(z) = \|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i| + \epsilon$ and the upper bounds of the error terms $\mathbf{E}_1$ and $\mathbf{E}_2$ can be found as

$$\begin{aligned} \|\mathbf{E}_1\| &\leq \frac{\max(z)}{\Delta - \max(z)} \sum_{i=1}^m |\alpha_i| \leq \frac{\max(z)}{\Delta - \max(z)} \left(\sum_{i=1}^m |\alpha_i|^{1/2} \right)^2 \\ \|\mathbf{E}_2\| &\leq \frac{\max(z)}{\Delta - \max(z)} \left(\sum_{i=1}^m |\alpha_i|^{1/2} \right)^2 . \end{aligned} \tag{2.17}$$

From the definition in Equation 2.12 we see that $\|\mathbf{G}_+(z)\| \leq \frac{1}{\Delta - \max(z)}$. Hence the norm of E_3 can be bounded by

$$\begin{aligned} \|\mathbf{E}_3\| &\leq \sum_{k=1}^{\infty} \frac{\| \sum_{i=1}^m (\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i)\|^2 \cdot \|\mathbf{H}_{\text{else}}\| + \frac{1}{\Delta} \sum_{i=1}^m (\kappa_i^2 \mathbf{A}_i^2 + \lambda_i^2 \mathbf{B}_i^2)\mathbf{I}\|^k}{(\Delta - \max(z))^{k+1}} \\ &\leq \sum_{k=1}^{\infty} \frac{2\Delta (\sum_{i=1}^m |\alpha_i|^{1/2})^2 (\|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|)^k}{(\Delta - \max(z))^{k+1}} \\ &= \frac{2\Delta (\sum_{i=1}^m |\alpha_i|^{1/2})^2}{\Delta - \max(z)} \cdot \frac{\|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|}{\Delta - \max(z) - (\|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|)}. \end{aligned} \tag{2.18}$$

Similar to the discussion in Section 2.2, to ensure that $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{targ}} \otimes \mathbf{P}_-\| \leq \epsilon$, which is the condition of Theorem 1.3.1, it is sufficient to let $\|\mathbf{E}_1\| + \|\mathbf{E}_2\| + \|\mathbf{E}_3\| \leq \epsilon$:

$$\begin{aligned} \|\mathbf{E}_1\| + \|\mathbf{E}_2\| + \|\mathbf{E}_3\| &\leq \frac{2\max(z)}{\Delta - \max(z)} \left(\sum_{i=1}^m |\alpha_i|^{1/2} \right)^2 \\ &+ \frac{2\Delta (\sum_{i=1}^m |\alpha_i|^{1/2})^2}{\Delta - \max(z)} \cdot \frac{\|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|}{\Delta - \max(z) - (\|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|)} \\ &= \frac{2(\sum_{i=1}^m |\alpha_i|^{1/2})^2 (\max(z) + \|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|)}{\Delta - \max(z) - (\|\mathbf{H}_{\text{else}}\| + \sum_{i=1}^m |\alpha_i|)} \leq \epsilon \end{aligned} \tag{2.19}$$

where we find the lower bound of Δ for parallel subdivision gadget

$$\Delta \geq \left[\frac{2(\sum_{i=1}^m |\alpha_i|^{1/2})^2}{\epsilon} + 1 \right] (2\|\mathbf{H}_{\text{else}}\| + 2\sum_{i=1}^m |\alpha_i| + \epsilon). \tag{2.20}$$

Note that if one substitutes $m = 1$ into Equation 2.20 the resulting expression is a lower bound that is less tight than that in Equation 2.9. This is because of the difference in the perturbation V between Equation 2.13 and Equation 3.5 which is explained in the text preceding Equation 2.14. Also we observe that the scaling of this lower bound for Δ is $O(\text{poly}(m)/\epsilon)$ for m parallel applications of subdivision gadgets, assuming $|\alpha_i| = O(\text{poly}(m))$ for every $i \in \{1, 2, \cdots, m\}$. This confirms the statement in [7, 8, 129] that subdivision gadgets can be applied to multiple terms in parallel and the scaling of the gap Δ in the case of m parallel subdivision gadgets will only differ to that of a single subdivision gadget by a polynomial in m.

Iterative scheme for k- to 3-body reduction. The iterative scheme in Algorithm 1 summarizes how to use parallel subdivision gadgets for reducing a k-body Ising Hamiltonian to 3-body (Here we use superscript $^{(i)}$ to represent the i^{th} iteration and subscript $_i$ for labelling objects within the same iteration).

We could show that after s iterations, the maximum spectral error between $\mathbf{\Pi}_-^{(s)}\widetilde{\mathbf{H}}^{(s)}\mathbf{\Pi}_-^{(s)}$ and $\widetilde{\mathbf{H}}^{(0)}\bigotimes_{i=1}^{s}\mathbf{P}_-^{(s)}$ is guaranteed to be within $s\epsilon$. Suppose we would like to make target Hamiltonian $\widetilde{\mathbf{H}}_0$, we construct a gadget $\widetilde{\mathbf{H}} = \mathbf{H}^{(1)} + \mathbf{V}^{(1)}$ according to Algorithm (1), such that $|\lambda(\widetilde{\mathbf{H}}^{(1)}) - \lambda(\widetilde{\mathbf{H}}^{(0)})| \leq \epsilon$ for low-lying eigenvalues $\lambda(\cdot)$. Note that in a precise sense we should write $|\lambda(\mathbf{\Pi}_-^{(1)}\widetilde{\mathbf{H}}^{(1)}\mathbf{\Pi}_-^{(1)}) - \lambda(\widetilde{\mathbf{H}}^{(0)} \otimes \mathbf{P}_-^{(0)})|$. Since the projectors $\mathbf{\Pi}_-^{(i)}$ and $\mathbf{P}_-^{(i)}$ do not affect the low-lying spectrum of $\widetilde{\mathbf{H}}^{(i)}$ and $\widetilde{\mathbf{H}}^{(i-1)}$, for simplicity and clarity we write only $\widetilde{\mathbf{H}}^{(i-1)}$ and $\widetilde{\mathbf{H}}^{(i)}$. After $\widetilde{\mathbf{H}}^{(1)}$ is introduced, according to Algorithm (1) the second gadget $\widetilde{\mathbf{H}}^{(2)}$ is then constructed by considering the *entire* $\widetilde{\mathbf{H}}^{(1)}$ as the new target Hamiltonian and introducing ancilla particles with unperturbed Hamiltonian $\mathbf{H}^{(2)}$ and perturbation $\mathbf{V}^{(2)}$ such that the low-energy spectrum of $\widetilde{\mathbf{H}}^{(2)}$ approximates the spectrum of $\widetilde{\mathbf{H}}^{(1)}$ up to error ϵ. In other words $|\lambda(\widetilde{\mathbf{H}}^{(1)}) - \lambda(\widetilde{\mathbf{H}}^{(2)})| \leq \epsilon$. With the serial application of gadgets we have produced a sequence of Hamiltonians $\widetilde{\mathbf{H}}^{(0)} \to \widetilde{\mathbf{H}}^{(1)} \to \widetilde{\mathbf{H}}^{(2)} \to \cdots \to \widetilde{\mathbf{H}}^{(k)}$ where $\widetilde{\mathbf{H}}^{(0)}$ is the target Hamiltonian and each subsequent gadget Hamiltonian $\widetilde{\mathbf{H}}^{(i)}$ captures the *entire* previous gadget $\widetilde{\mathbf{H}}^{(i-1)}$ in its low-energy sector with $|\lambda(\widetilde{\mathbf{H}}^{(i)}) - \lambda(\widetilde{\mathbf{H}}^{(i-1)})| \leq \epsilon$. Hence

Algorithm 1: Iterative scheme for reducing k-body Hamiltonian to 2-body

---

$\widetilde{\mathbf{H}}^{(0)} = \mathbf{H}_{\text{targ}}$; $\mathbf{H}_{\text{targ}}$ acts on the Hilbert space $\mathcal{M}^{(0)}$.

while $\widetilde{\mathbf{H}}^{(i)}$ is more than 3-body

Step 1: Find all the terms that are no more than 3-body (including $\mathbf{H}_{\text{else}}$ from $\widetilde{\mathbf{H}}^{(0)}$) in $\widetilde{\mathbf{H}}^{(i-1)}$ and let their sum be $\mathbf{H}^{(i)}_{\text{else}}$.

Step 2: Partition the rest of the terms in $\widetilde{\mathbf{H}}^{(i-1)}$ into $\alpha_1^{(i)}\mathbf{A}_1^{(i)} \otimes \mathbf{B}_1^{(i)}$, $\alpha_2^{(i)}\mathbf{A}_2^{(i)} \otimes \mathbf{B}_2^{(i)}$, $\cdots$, $\alpha_m^{(i)}\mathbf{A}_m^{(i)} \otimes \mathbf{B}_m^{(i)}$. Here $\alpha_j^{(i)}$ are coefficients.

Step 3: Introduce m ancilla qubits $w_1^{(i)}$, $w_2^{(i)}$, $\cdots w_m^{(i)}$ and construct $\widetilde{\mathbf{H}}^{(i)}$ using the parallel subdivision gadget.

Let $\mathbf{P}_-^{(i)} = |0\cdots 0\rangle\langle 0\cdots 0|_{w_1^{(i)}\cdots w_m^{(i)}}$ and $\mathbf{\Pi}_-^{(i)} = \mathbf{I}_{\mathcal{M}^{(i)}} \otimes \mathbf{P}_-^{(i)}$.

3.1: Apply the penalty Hamiltonian $\mathbf{H}^{(i)} = \sum_{x\in\{0,1\}^m} h(x)\Delta^{(i)}|x\rangle\langle x|$. Here $\Delta^{(i)}$ is calculated by the lower bound in Equation 2.20.

3.2: Apply the perturbation

$$\mathbf{V}^{(i)} = \mathbf{H}^{(i)}_{\text{else}} + \sum_{j=1}^m \sqrt{\tfrac{|\alpha_j^{(i)}|\Delta^{(i)}}{2}}(\text{sgn}(\alpha_j^{(i)})\mathbf{A}_j^{(i)} - \mathbf{B}_j^{(i)}) \otimes \mathbf{X}_{w_j^{(i)}} + \sum_{j=1}^m |\alpha_j^{(i)}|\mathbf{I}.$$

3.3: $\widetilde{\mathbf{H}}^{(i)} = \mathbf{H}^{(i)} + \mathbf{V}^{(i)}$ acts on the space $\mathcal{M}^{(i)}$ and the maximum spectral difference between $\widetilde{\mathbf{H}}_-^{(i)} = \mathbf{\Pi}_-^{(i)}\widetilde{\mathbf{H}}^{(i)}\mathbf{\Pi}_-^{(i)}$ and $\widetilde{\mathbf{H}}^{(i-1)} \otimes \mathbf{P}_-^{(i)}$ is at most ϵ.

Step 4: $i \leftarrow i + 1$.

end

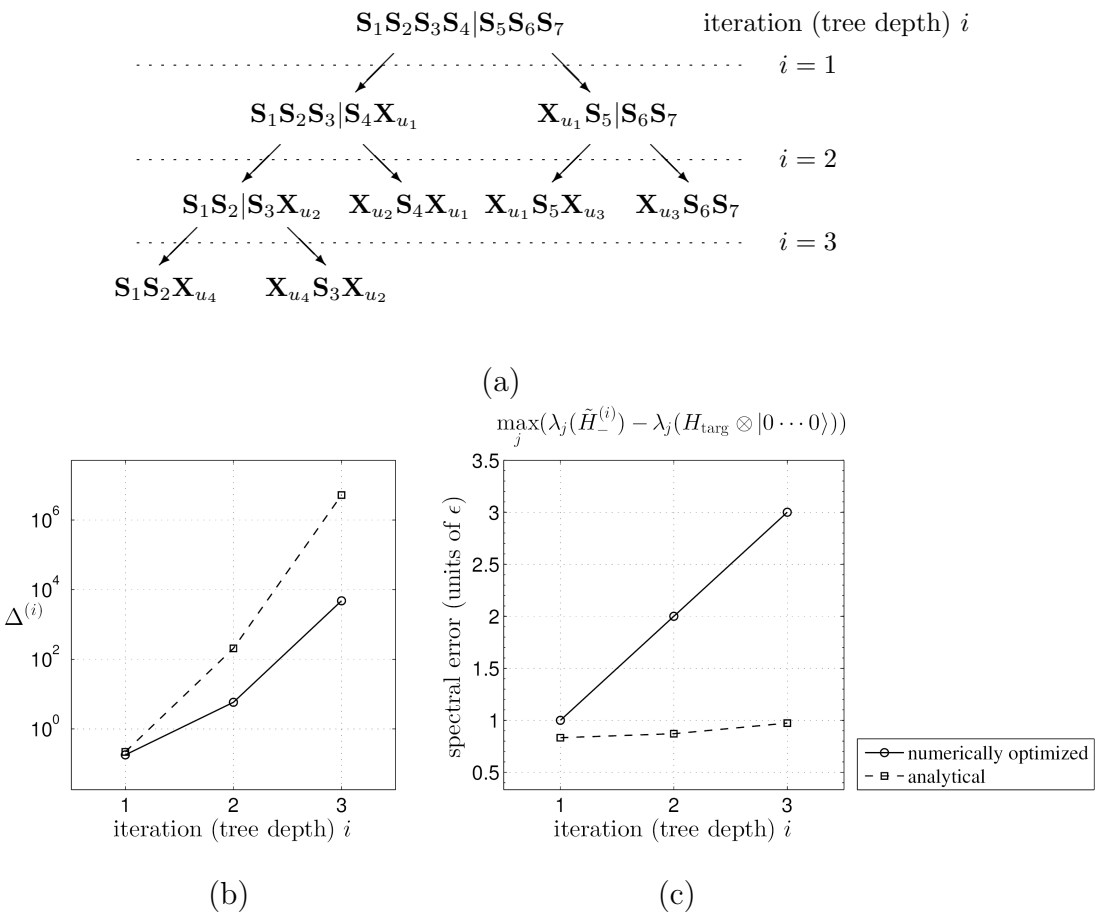


Fig. 2.3. (a): Reduction tree diagram for reducing a 7-body term to 3-body using parallel subdivision gadgets. Each $\mathbf{S}_i$ is a single-qubit Pauli operator acting on qubit i. The vertical lines $|$ show where the subdivisions are made at each iteration to each term. (b): An example where we consider the target Hamiltonian $\mathbf{H}_{\text{targ}} = \alpha\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3\mathbf{S}_4\mathbf{S}_5\mathbf{S}_6\mathbf{S}_7$ with $\alpha = 5\times10^{-3}$, $\mathbf{S}_i = \mathbf{X}_i$, $\forall i \in \{1, 2, \cdots, 7\}$, and reduce it to 3-body according to (a) up to error $\epsilon = 5\times10^{-4}$. This plot shows the energy gap applied onto the ancilla qubits introduced at each iteration. (c): The spectral error between the gadget Hamiltonian at each iteration $\widetilde{\mathbf{H}}^{(i)}$ and the target Hamiltonian $\mathbf{H}_{\text{targ}}$. For both (b)(c) the data labelled as "numerical" correspond to the case where during each iteration $\Delta^{(i)}$ is optimized such that the maximum spectral difference between $\mathbf{\Pi}_-^{(i)}\widetilde{\mathbf{H}}^{(i)}\mathbf{\Pi}_-^{(i)}$ and $\widetilde{\mathbf{H}}^{(i-1)} \otimes \mathbf{P}_-^{(i)}$ is ϵ. For definitions of $\Delta^{(i)}$, $\widetilde{\mathbf{H}}^{(i)}$, $\mathbf{\Pi}_-^{(i)}$ and $\mathbf{P}_-^{(i)}$, see Algorithm 1. Those labelled as "analytical" correspond to cases where each iteration uses the gap bound derived in Equation 2.20.

to bound the spectral error between the last gadget $\widetilde{\mathbf{H}}^{(k)}$ and the target Hamiltonian $\widetilde{\mathbf{H}}^{(0)}$ we could use triangle inequality: $|\lambda(\widetilde{\mathbf{H}}^{(s)}) - \lambda(\widetilde{\mathbf{H}}^{(0)})| \leq |\lambda(\widetilde{\mathbf{H}}^{(s)}) - \lambda(\widetilde{\mathbf{H}}^{(s-1)})| + \cdots + |\lambda(\widetilde{\mathbf{H}}^{(1)}) - \lambda(\widetilde{\mathbf{H}}^{(0)})| \leq s\epsilon$.

Total number of iterations for a k- to 3-body reduction. In general, given a k-body Hamiltonian, we apply the following parallel reduction scheme at each iteration until every term is 3-body: if k is even, this reduces it to two $(k/2+1)$-body terms; if k is odd, this reduces it to a $(\frac{k+1}{2}+1)$- and a $(\frac{k-1}{2}+1)$-body term. Define a function f such that a k-body term needs $f(k)$ iterations to be reduced to 3-body. Then we have the recurrence

$$f(k) = \begin{cases} f\left(\frac{k}{2}+1\right)+1 & k \text{ even} \\ f\left(\frac{k+1}{2}+1\right)+1 & k \text{ odd} \end{cases} \tag{2.21}$$

with $f(3) = 0$ and $f(4) = 1$. One can check that $f(k) = \lceil \log_2(k-2) \rceil$, $k \geq 4$ satisfies this recurrence. Therefore, using subdivision gadgets, one can reduce a k-body interaction to 3-body in $s = \lceil \log_2(k-2) \rceil$ iterations and the spectral error between $\widetilde{\mathbf{H}}^{(s)}$ and $\widetilde{\mathbf{H}}^{(0)}$ is within $\lceil \log_2(k-2) \rceil \epsilon$.

Gap scaling. From the iterative scheme shown previously one can conclude that $\Delta^{(i+1)} = \Theta(\epsilon^{-1}(\Delta^{(i)})^{3/2})$ for the $(i+1)^{\text{th}}$ iteration, which implies that for a total of s iterations,

$$\Delta^{(s)} = \Theta\left(\epsilon^{-2[(3/2)^{s-1}-1]}(\Delta^{(1)})^{(3/2)^{s-1}}\right). \tag{2.22}$$

Since $s = \lceil \log_2(k-2) \rceil$ and $\Delta^{(1)} = \Theta(\epsilon^{-1})$ we have

$$\Delta^{(s)} = \Theta\left(\epsilon^{-3(\frac{1}{2}\lceil k-2 \rceil)^{\log_2(3/2)}-2}\right) = \Theta\left(\epsilon^{-\text{poly}(k)}\right) \tag{2.23}$$

accumulating exponentially as a function of k. The exponential nature of the scaling with respect to k agrees with results by Bravyi et al. [129]. However, in our con-

struction, due to the improvement of gap scaling in a single subdivision gadget from $\Delta = \Theta(\epsilon^{-2})$ to $\Theta(\epsilon^{-1})$, the scaling exponents in

$$\Delta^{(i+1)} = \Theta(\epsilon^{-1}(\Delta^{(i)})^{3/2})$$

are also improved quadratically over those in [129], which is $\Delta^{(i+1)} = \Theta(\epsilon^{-2}(\Delta^{(i)})^3)$.

Qubit cost. Based on the reduction scheme described in Algorithm 1 (illustrated in Figure 2.3a for 7-body), the number of ancilla qubits needed for reducing a k-body term to 3-body is $k-3$. Suppose we are given a k-body target term $\mathbf{S}_1\mathbf{S}_2\cdots\mathbf{S}_k$ (where all of the operators S_i act on separate spaces) and we would like to reduce it to 3-body using the iterative scheme in Algorithm 1. At each iteration, if we describe every individual subdivision gadget by a vertical line | at the location where the partition is made, for example $\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3\mathbf{S}_4|\mathbf{S}_5\mathbf{S}_6\mathbf{S}_7$ in the case of the first iteration in Figure 2.3a, then after $\lceil \log_2(k-2) \rceil$ iterations all the partitions made to the k-body term can be described as $\mathbf{S}_1\mathbf{S}_2|\mathbf{S}_3|\mathbf{S}_4|\cdots|\mathbf{S}_{k-2}|\mathbf{S}_{k-1}\mathbf{S}_k$. Note that there are $k-3$ vertical lines in total, each corresponding to an ancilla qubit needed for a subdivision gadget. Therefore in total $k-3$ ancilla qubits are needed for reducing a k-body term to 3-body.

Example: Reducing 7-body to 3-body. We have used numerics to test the reduction algorithm in Algorithm 1 on a target Hamiltonian $\mathbf{H}_{\text{targ}} = \alpha\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3\mathbf{S}_4\mathbf{S}_5\mathbf{S}_6\mathbf{S}_7$. Here we let $\mathbf{S}_i = \mathbf{X}_i$, $\forall i \in \{1, 2, \cdots, 7\}$, $\epsilon = 5 \times 10^{-4}$ and $\alpha = 5 \times 10^{-3}$. During each iteration the values of $\Delta^{(i)}$ are assigned according to the lower bound in Equation 2.20. From Figure 2.3c we can see that the lower bounds are sufficient for keeping the total spectral error between $\widetilde{\mathbf{H}}^{(3)}_{<E_*}$ and $\widetilde{\mathbf{H}}^{(0)} \bigotimes_{i=1}^{3} \mathbf{P}_-^{(i)}$ within 3ϵ. Furthermore, numerical search is also used at each iteration to find the minimum value of $\Delta^{(i)}$ so that the spectral error between $\mathbf{\Pi}_-^{(i)}\widetilde{\mathbf{H}}^{(i)}\mathbf{\Pi}_-^{(i)}$ and $\widetilde{\mathbf{H}}^{(i-1)} \bigotimes_{j=1}^{i} \mathbf{P}_-^{(j)}$ is ϵ. The numerically found gaps $\Delta^{(i)}$ are much smaller than their analytical counterparts at each iteration (Figure 2.3b), at the price that the error is larger (Figure 2.3c). In both the

numerical and the analytical cases, the error appears to accumulate linearly as the iteration proceeds.

2.4 Improved 3- to 2-body gadget

Summary. Subdivision gadgets cannot be used for reducing from 3- to 2-body; accordingly, the final reduction requires a different type of gadget [7,8,129]. Consider 3-body target Hamiltonian of the form $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$. Here $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are unit-norm Hermitian operators acting on separate spaces $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$. Here we focus on the gadget construction introduced in Oliveira and Terhal [8] and also used in Bravyi, DiVincenzo, Loss and Terhal [129]. To accomplish the 3- to 2-body reduction, we introduce an ancilla spin w and apply a penalty Hamiltonian $\mathbf{H} = \Delta |1\rangle\langle 1|_w$. We then add a perturbation $\mathbf{V}$ of form,

$$\mathbf{V} = \mathbf{H}_{\text{else}} + \mu \mathbf{C} \otimes |1\rangle\langle 1|_w + (\kappa \mathbf{A} + \lambda \mathbf{B}) \otimes \mathbf{X}_w + \mathbf{V}_1 + \mathbf{V}_2 \tag{2.24}$$

where $\mathbf{V}_1$ and $\mathbf{V}_2$ are 2-local compensation terms (details presented later in this section):

$$\begin{aligned} \mathbf{V}_1 &= \frac{1}{\Delta}(\kappa^2 + \lambda^2)|0\rangle\langle 0|_w + \frac{2\kappa\lambda}{\Delta}\mathbf{A} \otimes \mathbf{B} - \frac{1}{\Delta^2}(\kappa^2 + \lambda^2)\mu \mathbf{C} \otimes |0\rangle\langle 0|_w \\ \mathbf{V}_2 &= -\frac{2\kappa\lambda}{\Delta^3}\text{sgn}(\alpha)\Big[(\kappa^2 + \lambda^2)|0\rangle\langle 0|_w + 2\kappa\lambda \mathbf{A} \otimes \mathbf{B}\Big]. \end{aligned} \tag{2.25}$$

Here we let $\kappa = \text{sgn}(\alpha)\,(\alpha/2)^{1/3}\,\Delta^{3/4}$, $\lambda = (\alpha/2)^{1/3}\,\Delta^{3/4}$ and $\mu = (\alpha/2)^{1/3}\,\Delta^{1/2}$.

For sufficiently large Δ, the low-lying spectrum of the gadget Hamiltonian $\widetilde{\mathbf{H}}$ captures the entire spectrum of $\mathbf{H}_{\text{targ}}$ up to arbitrary error ϵ. In the construction of [129] it is shown that $\Delta = \Theta(\epsilon^{-3})$ is sufficient. In [7], $\Delta = \Theta(\epsilon^{-3})$ is also assumed, though the construction of $\mathbf{V}$ is slightly different from Equation 2.24. By adding terms in $\mathbf{V}$ to compensate for the perturbative error due to the modification, we find that $\Delta = \Theta(\epsilon^{-2})$ is sufficient for accomplishing the 3- to 2-body reduction:

$$\Delta \geq \frac{1}{4}(-b + \sqrt{b^2 - 4c})^2 \tag{2.26}$$

where b and c are defined as

$$\begin{aligned} b &= -\left[\xi + \frac{2^{4/3}\alpha^{2/3}}{\epsilon}(\max z + \eta + \xi^2)\right] \\ c &= -\left(1 + \frac{2^{4/3}\alpha^{2/3}}{\epsilon}\xi\right)(\max z + \eta) \end{aligned} \tag{2.27}$$

with $\max z = \|\mathbf{H}_{\text{else}}\| + |\alpha| + \epsilon$, $\eta = \|\mathbf{H}_{\text{else}}\| + 2^{2/3}\alpha^{4/3}$ and $\xi = 2^{-1/3}\alpha^{1/3} + 2^{1/3}\alpha^{2/3}$. From Equation 2.26 we can see the lower bound to Δ is $\Theta(\epsilon^{-2})$. Our improvement results in a power of ϵ^{-1} reduction in the gap. For the dependence of Δ on $\|\mathbf{H}_{\text{else}}\|$, α and ϵ^{-1} for both the original [8] and the optimized case, see Figure 2.4. Results show that the bound in Equation 2.26 is tight with respect to the minimum Δ numerically found that yields the spectral error between $\widetilde{\mathbf{H}}_{<E_*}$ and $\mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ to be ϵ.

Analysis. We will proceed by first presenting the improved construction of the 3- to 2-body gadget and then show that $\Delta = \Theta(\epsilon^{-2})$ is sufficient for the spectral error to be $\leq \epsilon$. Then we present the construction in the literature [8, 129] and argue that $\Delta = \Theta(\epsilon^{-3})$ is required for yielding a spectral error between $\widetilde{\mathbf{H}}$ and $\mathbf{H}_{\text{eff}}$ within ϵ using this construction.

In the improved construction we define the perturbation $\mathbf{V}$ as in Equation 2.24. Here the coefficients are chosen to be $\kappa = \Theta(\Delta^{3/4})$, $\lambda = \Theta(\Delta^{3/4})$ and $\mu = \Theta(\Delta^{1/2})$. In order to show that the assigned powers of Δ in the coefficients are optimal, we introduce a parameter r such that

$$\kappa = \text{sgn}(\alpha)\left(\frac{\alpha}{2}\right)^{1/3}\Delta^r, \qquad \lambda = \left(\frac{\alpha}{2}\right)^{1/3}\Delta^r, \qquad \mu = \left(\frac{\alpha}{2}\right)^{1/3}\Delta^{2-2r}. \tag{2.28}$$

It is required that $\|\mathbf{V}\| \leq \Delta/2$ (Theorem 1.3.1) for the convergence of the perturbative series. Therefore let $r < 1$ and $2 - 2r < 1$, which gives $1/2 < r < 1$. With the definitions $\mathcal{L}_-$ and $\mathcal{L}_+$ being the ground and excited state subspaces respectively, $\mathbf{V}_-$, $\mathbf{V}_+$, $\mathbf{V}_{-+}$, $\mathbf{V}_{+-}$ can be calculated as the following:

$$\begin{aligned} \mathbf{V}_- = \Bigg[\mathbf{H}_{\text{else}} + \frac{1}{\Delta}(\kappa\mathbf{A} + \lambda\mathbf{B})^2 - \frac{1}{\Delta}(\kappa^2 + \lambda^2)\mu\mathbf{C} \\ - \frac{2\kappa\lambda}{\Delta^3}\text{sgn}(\alpha)(\kappa\mathbf{A} + \lambda\mathbf{B})^2\Bigg] \otimes |0\rangle\langle 0|_w \end{aligned} \tag{2.29}$$

$$\mathbf{V}_+ = \left[\mathbf{H}_{\text{else}} + \mu\mathbf{C} + \frac{2\kappa\lambda}{\Delta}\mathbf{A}\otimes\mathbf{B} - \frac{4\kappa^2\lambda^2}{\Delta^3}\text{sgn}(\alpha)\mathbf{A}\otimes\mathbf{B}\right]\otimes|1\rangle\langle 1|_w \tag{2.30}$$

$$\mathbf{V}_{-+} = (\kappa\mathbf{A}+\lambda\mathbf{B})\otimes|0\rangle\langle 1|_w \tag{2.31}$$

$$\mathbf{V}_{+-} = (\kappa\mathbf{A}+\lambda\mathbf{B})\otimes|1\rangle\langle 0|_w. \tag{2.32}$$

The self-energy expansion, referring to Equation 1.26, becomes

$$\begin{aligned}
\mathbf{\Sigma}_-(z) &= \mathbf{V}_- + \frac{1}{z-\Delta}\mathbf{V}_{-+}\mathbf{V}_{+-} + \frac{1}{(z-\Delta)^2}\mathbf{V}_{-+}\mathbf{V}_+\mathbf{V}_{+-} + \sum_{k=2}^{\infty}\frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}} \\
&= \underbrace{\mathbf{H}_{\text{else}}}_{(a)} + \underbrace{\frac{1}{\Delta}(\kappa\mathbf{A}+\lambda\mathbf{B})^2}_{(b)}\underbrace{-\frac{1}{\Delta}(\kappa^2+\lambda^2)\mu\mathbf{C}}_{(c)}\underbrace{-\frac{2\kappa\lambda}{\Delta^3}\text{sgn}(\alpha)(\kappa\mathbf{A}+\lambda\mathbf{B})^2}_{(d)} \\
&+ \underbrace{\frac{1}{z-\Delta}(\kappa\mathbf{A}+\lambda\mathbf{B})^2}_{(e)} + \frac{1}{(z-\Delta)^2}(\kappa\mathbf{A}+\lambda\mathbf{B})\left[\underbrace{\mathbf{H}_{\text{else}}}_{(f)} + \underbrace{\mu\mathbf{C}}_{(g)}\right. \\
&+ \left.\underbrace{\frac{2\kappa\lambda}{\Delta}\mathbf{A}\otimes\mathbf{B}}_{(h)}\underbrace{-\frac{4\kappa^2\lambda^2}{\Delta^3}\text{sgn}(\alpha)\mathbf{A}\otimes\mathbf{B}}_{(i)}\right](\kappa\mathbf{A}+\lambda\mathbf{B}) + \underbrace{\sum_{k=2}^{\infty}\frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}}}_{(j)}.
\end{aligned} \tag{2.33}$$

Now we rearrange the terms in the self energy expansion so that the target Hamiltonian arising from the leading order terms can be separated from the rest, whcih are error terms. Observe that term (g) combined with the factors outside the bracket could give rise to a 3-body $\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}$ term:

$$\begin{aligned}
\frac{1}{(z-\Delta)^2}(\kappa\mathbf{A}+\lambda\mathbf{B})^2\mu\mathbf{C} &= \underbrace{\frac{2\kappa\lambda\mu}{\Delta^2}\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}}_{(g_1)} \\
&+ \underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)2\kappa\lambda\mu\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}}_{(g_2)} \\
&+ \underbrace{\frac{1}{(z-\Delta)^2}(\kappa^2+\lambda^2)\mu\mathbf{C}}_{(g_3)}.
\end{aligned} \tag{2.34}$$

Here (g_1) combined with term (a) in (2.33) gives $\mathbf{H}_{\text{targ}}$. (g_2) and (g_3) are error terms. Now we further rearrange the error terms as the following. We combine term (b) and (e) to form $\mathbf{E}_1$, term (c) and (g_3) to form $\mathbf{E}_2$, term (f) and the factors outside

the bracket to be $\mathbf{E}_3$. Rename (g_2) to be $\mathbf{E}_4$. Using the identity $(\kappa\mathbf{A}+\lambda\mathbf{B})(A\otimes B)(\kappa A+\lambda B) = \mathrm{sgn}(\alpha)(\kappa\mathbf{A}+\lambda\mathbf{B})^2$ we combine term (d) and (h) along with the factors outside the bracket to be $\mathbf{E}_5$. Rename (i) to be $\mathbf{E}_6$ and (j) to be $\mathbf{E}_7$. The rearranged self-energy expanision reads

$$\begin{aligned}
\Sigma_-(z) &= \Bigg[\underbrace{\mathbf{H}_{\text{else}}+\frac{2\kappa\lambda\mu}{\Delta^2}\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}}_{\mathbf{H}_{\text{targ}}}+\underbrace{\left(\frac{1}{\Delta}+\frac{1}{z-\Delta}\right)(\kappa\mathbf{A}+\lambda\mathbf{B})^2}_{\mathbf{E}_1}\\
&+\underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)(\kappa^2+\lambda^2)\mu\mathbf{C}}_{\mathbf{E}_2}+\underbrace{\frac{1}{(z-\Delta)^2}(\kappa\mathbf{A}+\lambda\mathbf{B})\mathbf{H}_{\text{else}}(\kappa\mathbf{A}+\lambda\mathbf{B})}_{\mathbf{E}_3}\\
&+\underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)2\kappa\lambda\mu\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}}_{\mathbf{E}_4}\\
&+\underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)\frac{2\kappa\lambda}{\Delta}\mathrm{sgn}(\alpha)(\kappa\mathbf{A}+\lambda\mathbf{B})^2}_{\mathbf{E}_5}\\
&-\underbrace{\frac{1}{(z-\Delta)^2}\cdot\frac{4\kappa^2\lambda^2}{\Delta^3}(\kappa\mathbf{A}+\lambda\mathbf{B})^2}_{\mathbf{E}_6}\Bigg]\otimes|0\rangle\langle 0|_w+\underbrace{\sum_{k=2}^{\infty}\frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}}}_{\mathbf{E}_7}.
\end{aligned} \tag{2.35}$$

We bound the norm of each error term in the self energy expansion Equation 2.35 by substituting the definitions of κ, λ and μ in Equation 2.28 and letting z be the maximum value permitted by Theorem 1.3.1 which is $\max z = |\alpha|+\epsilon+\|\mathbf{H}_{\text{else}}\|$:

$$\begin{aligned}
\|\mathbf{E}_1\| &\le \frac{\max z\cdot 2^{4/3}\alpha^{2/3}\Delta^{2r-1}}{\Delta-\max z} = \Theta(\Delta^{2r-2}),\\
\|\mathbf{E}_2\| &\le \frac{(2\Delta-\max z)\max z}{(\Delta-\max z)^2}\cdot\alpha = \Theta(\Delta^{-1}),
\end{aligned} \tag{2.36}$$

$$\begin{aligned}
\|\mathbf{E}_3\| &\le \frac{2^{4/3}\alpha^{2/3}\Delta^{2r}\|\mathbf{H}_{\text{else}}\|}{(\Delta-\max z)^2} = \Theta(\Delta^{2r-2}),\\
\|\mathbf{E}_4\| &\le \frac{(2\Delta-\max z)\max z}{(\Delta-\max z)^2}\cdot\alpha = \Theta(\Delta^{-1}),
\end{aligned} \tag{2.37}$$

$$\begin{aligned}
\|\mathbf{E}_5\| &\le \frac{(2\Delta-\max z)\max z}{(\Delta-\max z)^2}\cdot 2^{5/3}\alpha^{4/3}\Delta^{4r-3} = \Theta(\Delta^{4r-4}),\\
\|\mathbf{E}_6\| &\le \frac{4\alpha^2\Delta^{6r-3}}{(\Delta-\max z)^2} = \Theta(\Delta^{6r-5}),
\end{aligned} \tag{2.38}$$

$$\begin{aligned}
\|\mathbf{E}_7\| &\leq \sum_{k=2}^{\infty} \left\| \frac{(\kappa A + \lambda B)\left(\mathbf{H}_{\text{else}} + \mu\mathbf{C} + \frac{2\kappa\lambda}{\Delta}\left(1 + \frac{2\kappa\lambda}{\Delta^2}\right)\mathbf{A}\otimes\mathbf{B}\right)^k (\kappa\mathbf{A} + \lambda\mathbf{B})}{(\Delta - \max z)^{k+1}} \right\| \\
&\leq \frac{2^{4/3}\alpha^{2/3}\Delta^{2r}}{(\Delta - \max z)} \sum_{k=2}^{\infty} \frac{1}{(\Delta - \max z)^k} \Big(\|\mathbf{H}_{\text{else}}\| + 2^{-1/3}\alpha^{1/3}\Delta^{2-2r} \\
&\qquad + 2^{1/3}\alpha^{2/3}\Delta^{2r-1} + 2^{2/3}\alpha^{4/3}\Delta^{4r-3} \Big)^k \\
&= \Theta(\Delta^{\max\{1-2r, 6r-5, 10r-9\}}).
\end{aligned} \tag{2.39}$$

Now the self energy expansion can be written as

$$\mathbf{\Sigma}_-(z) = \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w + \Theta(\Delta^{f(r)})$$

where the function $f(r) < 0$ determines the dominant power in Δ from $\|\mathbf{E}_1\|$ through $\|\mathbf{E}_6\|$:

$$f(r) = \max\{1 - 2r, 6r - 5\}, \quad \frac{1}{2} < r < 1. \tag{2.40}$$

In order to keep the error $O(\epsilon)$, it is required that $\Delta = \Theta(\epsilon^{1/f(r)})$. To optimize the gap scaling as a function of ϵ, $f(r)$ must take the minimum value. As is shown in Figure 2.5b, when $r = 3/4$, the minimum value $f(r) = -1/2$ is obtained, which corresponds to $\Delta = \Theta(\epsilon^{-2})$. We have hence shown that the powers of Δ in the assignments of κ, λ and μ in Equation 2.28 are optimal for the improved gadget construction. The optimal scaling of $\Theta(\epsilon^{-2})$ is also numerically confirmed in Figure 2.4a. As one can see, the optimized slope $\mathrm{d}\log\Delta/\mathrm{d}\log\epsilon^{-1}$ is approximately 2 for small ϵ.

One natural question to ask next is whether it is possible to further improve the gap scaling as a function of ϵ. This turns out to be difficult. Observe that the $6r - 5$ component of $f(r)$ in Equation 2.40 comes from E_6 and E_7 in Equation 2.35. In E_7, the $\Theta(\Delta^{6r-5})$ contribution is attributed to the term $\frac{1}{\Delta}(\kappa\mathbf{A} + \lambda\mathbf{B})^2$ in V_1 of Equation 2.25, which is intended for compensating the 2^{nd} order perturbative term and therefore cannot be removed from the construction.

We now let $r = 3/4$ be a fixed constant and derive the lower bound for Δ such that for given α, $\mathbf{H}_{\text{else}}$ and ϵ, the spectral error between the effective Hamiltonian

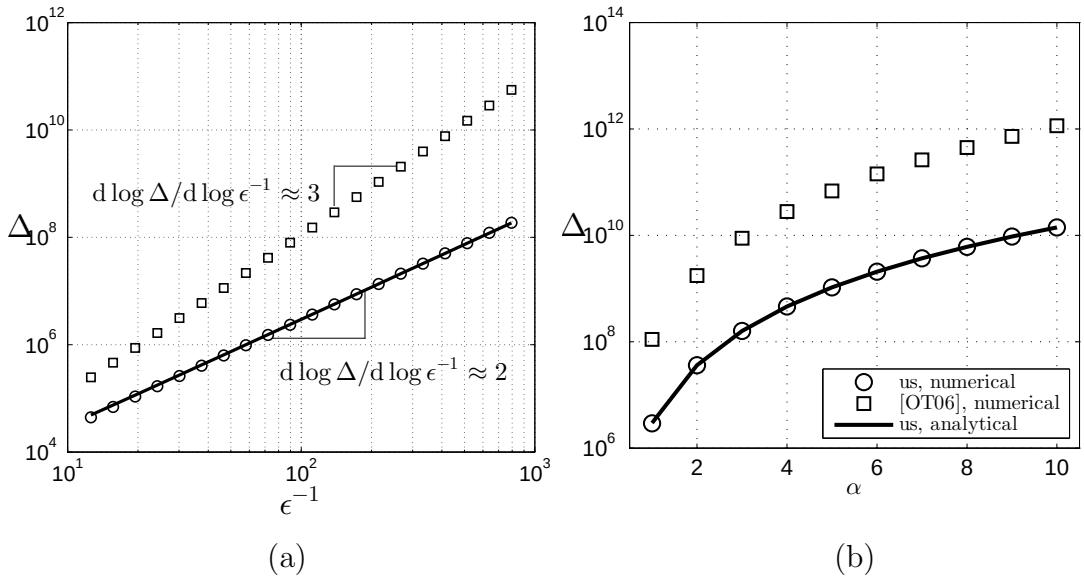


Fig. 2.4. Comparison between our 3- to 2-body gadget with that of Oliveira and Terhal [8]. As Δ is not explicitly assigned as a function of α, $\|\mathbf{H}_{\text{else}}\|$ and ϵ in [8], we numerically find the optimal Δ values for their constructions (marked as "[OT06]"). Subplot (a) shows the scaling of the gap Δ as a function of error tolerance ϵ. Subplot (b) shows the gap Δ as a function of the desired coupling α. For the meanings of the labels in the legend, see Figure 2.2. The fixed parameters in each subplots are: (a) $\|\mathbf{H}_{\text{else}}\| = 0$, $\alpha = 1$. (b) $\epsilon = 0.01$, $\|\mathbf{H}_{\text{else}}\| = 0$. Note that our constructions have improved the Δ scaling for the ranges of α and ϵ considered.

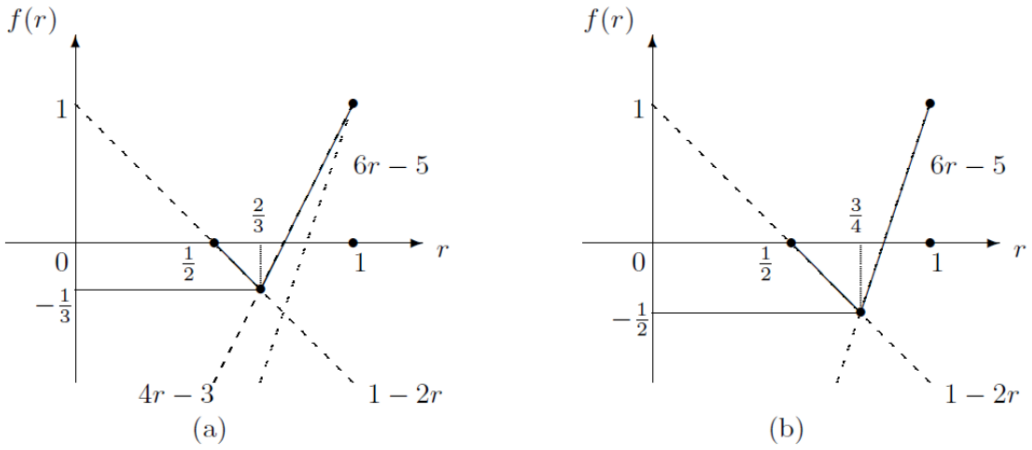


Fig. 2.5. The function $f(r)$ shows the dominant power of Δ in the error terms in the perturbative expansion. (a): When the error term $\mathbf{E}_4$ in Equation 2.48, which contributes to the $4r-3$ component of $f(r)$ in Equation 2.50, is not compensated in the original construction by Oliveira and Terhal, the dominant power of Δ in the error term $f(r)$ takes minimum value of $-1/3$, indicating that $\Delta = \Theta(\epsilon^{-3})$ is required. (b): In the improved construction, $\min_{r\in(1/2,1)} f(r) = -1/2$ indicating that $\Delta = \Theta(\epsilon^{-2})$.

$\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ and $\widetilde{\mathbf{H}}_{<E_*}$ is within ϵ. This amounts to satisfying the condition of Theorem 1.3.1:

$$\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\| \leq \epsilon. \tag{2.41}$$

Define the total error $\mathbf{E} = \mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}} = \mathbf{E}_1 + \cdots + \mathbf{E}_7$. For convenience we also define $\eta = \|\mathbf{H}_{\text{else}}\| + 2^{2/3}\alpha^{4/3}$ and $\xi = 2^{-1/3}\alpha^{1/3} + 2^{1/3}\alpha^{2/3}$. Then

$$\begin{aligned} \|E_7\| &\leq \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}}{\Delta - \max z} \sum_{k=2}^{\infty} \frac{(\eta + \xi\Delta^{1/2})^k}{(\Delta - \max z)^k} \\ &= \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}}{\Delta - \max z - (\eta + \xi\Delta^{1/2})} \left(\frac{\eta + \xi\Delta^{1/2}}{\Delta - \max z}\right)^2. \end{aligned} \tag{2.42}$$

The upper bound for $\|\mathbf{E}\|$ is then found by summing over Equation 2.36, 2.37, 2.38 and 2.42:

$$\begin{aligned} \|\mathbf{E}\| &\leq \frac{\max z \cdot 2^{4/3}\alpha^{2/3}\Delta^{1/2}}{\Delta - \max z} + \frac{(2\Delta - \max z)\max z}{(\Delta - \max z)^2} \cdot 2^{4/3}\alpha^{3/2}\xi + \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}\eta}{(\Delta - \max z)^2} \\ &+ \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}}{\Delta - \max z - (\eta + \xi\Delta^{1/2})} \left(\frac{\eta + \xi\Delta^{1/2}}{\Delta - \max z}\right)^2. \end{aligned} \tag{2.43}$$

By rearranging the terms in Equation 2.43 we arrive at a simplified expression for the upper bound presented below. Requiring the upper bound of $\|\mathbf{E}\|$ to be within ϵ gives

$$\|\mathbf{E}\| \leq 2^{4/3}\alpha^{2/3} \frac{(\max z + \eta + \xi^2)\Delta^{1/2} + \xi(\max z + \eta)}{\Delta - \xi\Delta^{1/2} - (\max z + \eta)} \leq \epsilon. \tag{2.44}$$

Equation 2.44 is a quadratic constraint with respect to $\Delta^{1/2}$. Solving the inequality gives the lower bound of Δ given in Equation 2.26. Note here that $\Delta = \Theta(\epsilon^{-2})$, which improves over the previously assumed $\Delta = \Theta(\epsilon^{-3})$ in the literature [7, 8, 129]. This bound is shown in Figure 2.4b as the "analytical lower bound". Comparison between the analytical lower bound and the numerically optimized gap in Figure 2.4b indicates that the lower bound is relatively tight when $\|\mathbf{H}_{\text{else}}\| = 0$. If $\mathbf{H}_{\text{else}}$ is non-zero, the bound is likely going to be less tight because $\mathbf{H}_{\text{else}}$ may not commute with the other terms in the target Hamiltonian.

Comparison with Oliveira and Terhal [8]. Given operators $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{T}$ acting on separate spaces $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ respectively, the 3- to 2-body construction in [7, 8] approximates the target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \mathbf{Q} \otimes \mathbf{R} \otimes \mathbf{T}$. In order to compare with their construction, however, we let $\alpha = \|\mathbf{Q}\| \cdot \|\mathbf{R}\| \cdot \|\mathbf{T}\|$ and define $\mathbf{Q} = \alpha^{1/3}\mathbf{A}$, $\mathbf{R} = \alpha^{1/3}\mathbf{B}$ and $\mathbf{T} = \alpha^{1/3}\mathbf{C}$. Hence the target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}$ with $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ being unit-norm Hermitian operators. Introduce an ancilla qubit w and apply the penalty Hamiltonian $\mathbf{H} = \Delta|1\rangle\langle 1|_w$. In the construction by Oliveira and Terhal [8], the perturbation $\mathbf{V}$ is defined as

$$\mathbf{V} = \mathbf{H}_{\text{else}} \otimes \mathbf{I}_w + \mu\mathbf{C} \otimes |1\rangle\langle 1|_w + (\kappa\mathbf{A} + \lambda\mathbf{B}) \otimes \mathbf{X}_w + \mathbf{V}_1' \tag{2.45}$$

where the compensation term $\mathbf{V}_1'$ is

$$\mathbf{V}_1' = \frac{1}{\Delta}(\kappa\mathbf{A} + \lambda\mathbf{B})^2 - \frac{1}{\Delta^2}(\kappa^2\mathbf{A}^2 + \lambda^2\mathbf{B}^2)\mu\mathbf{C}. \tag{2.46}$$

Comparing Equation 2.46 with the expression for $\mathbf{V}_1$ in Equation 2.25, one observes that $\mathbf{V}_1$ slightly improves over $\mathbf{V}_1'$ by projecting 1-local terms to $\mathcal{L}_-$ so that $\mathbf{V}$ will have less contribution to $\mathbf{V}_+$, which reduces the high order error terms in the perturbative expansion. However, this modification comes at a cost of requiring more 2-local terms in the perturbation $\mathbf{V}$.

From the gadget construction shown in [8, Equation 26], the equivalent choices of the coefficients κ, λ and μ are

$$\kappa = -\left(\frac{\alpha}{2}\right)^{1/3}\frac{1}{\sqrt{2}}\Delta^r, \quad \lambda = \left(\frac{\alpha}{2}\right)^{1/3}\frac{1}{\sqrt{2}}\Delta^r, \quad \mu = -\left(\frac{\alpha}{2}\right)^{1/3}\Delta^{2-2r} \tag{2.47}$$

where $r = 2/3$ in the constructions used in [8, 129]. In fact this value of r is optimal for the construction in the sense that it leads to the optimal gap scaling $\Delta = \Theta(\epsilon^{-3})$. Expanding the self-energy to 3rd order, following a similar procedure as in (2.33), we have

$$\Sigma_-(z) = \Bigg[\underbrace{\mathbf{H}_{\text{else}} + \frac{2\kappa\lambda\mu}{\Delta^2}\mathbf{A}\otimes\mathbf{B}\otimes\mathbf{C}}_{\mathbf{H}_{\text{targ}}} + \underbrace{\left(\frac{1}{\Delta} + \frac{1}{z-\Delta}\right)(\kappa\mathbf{A} + \lambda\mathbf{B})^2}_{\mathbf{E}_1}$$

$$
\begin{aligned}
&+ \underbrace{\left(\frac{1}{(z-\Delta)^2} - \frac{1}{\Delta^2}\right)(\kappa^2\mathbf{A}^2 + \lambda^2\mathbf{B}^2)\mu\mathbf{C}}_{\mathbf{E}_2} \\
&+ \underbrace{\frac{1}{(z-\Delta)^2}(\kappa\mathbf{A} + \lambda\mathbf{B})\mathbf{H}_{\text{else}}(\kappa\mathbf{A} + \lambda\mathbf{B})}_{\mathbf{E}_3} \\
&+ \underbrace{\frac{1}{(z-\Delta)^2} \cdot \frac{1}{\Delta}(\kappa\mathbf{A} + \lambda\mathbf{B})^4}_{\mathbf{E}_4} \\
&- \underbrace{\frac{1}{(z-\Delta)^2} \cdot \frac{1}{\Delta^2}(\kappa^2\mathbf{A}^2 + \lambda^2\mathbf{B}^2)\mu(\kappa\mathbf{A} + \lambda\mathbf{B})^2\mathbf{C}}_{E_5}\Bigg] \otimes |0\rangle\langle 0|_w \\
&+ \underbrace{\sum_{k=2}^{\infty} \frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}}}_{\mathbf{E}_6}.
\end{aligned} \tag{2.48}
$$

Similar to the derivation of Equation 2.36, 2.37, and 2.38 by letting $z \mapsto \max z$, where $\max z = |\alpha| + \epsilon + \|\mathbf{H}_{\text{else}}\|$ is the largest value of z permitted by the Theorem 1.3.1, and using the triangle inequality to bound the norm, we can bound the norm of the error terms $\mathbf{E}_1$ through $\mathbf{E}_6$. For example,

$$\|\mathbf{E}_1\| \leq \left(\frac{1}{\Delta - \max z} - \frac{1}{\Delta}\right) 2^2 \cdot \left(\frac{\alpha}{2}\right)^{2/3} \Delta^{2r} = \Theta(\Delta^{2r-2}).$$

Applying the same calculation to $\mathbf{E}_2, \mathbf{E}_3, \cdots$ we find that $\|\mathbf{E}_2\| = \Theta(\Delta^{-1})$, $\|E_3\| = \Theta(\Delta^{2r-2})$, $\|\mathbf{E}_4\| = \Theta(\Delta^{4r-3})$, $\|\mathbf{E}_5\| = \Theta(\Delta^{4r-4})$. The norm of the high order terms $\mathbf{E}_6$ can be bounded as

$$
\begin{aligned}
\|\mathbf{E}_6\| &\leq \sum_{k=2}^{\infty} \frac{\|\mathbf{V}_{-+}\| \cdot \|\mathbf{V}_+\|^k \cdot \|\mathbf{V}_{+-}\|}{(\Delta - \max(z))^{k+1}} \leq \frac{4\left(\frac{\alpha}{2}\right)^{1/3}\Delta^{2r}}{\Delta - \max(z)} \sum_{k=2}^{\infty} \left(\frac{\rho}{\Delta - \max(z)}\right)^k \\
&= \frac{2^{4/3}\alpha^{2/3}\Delta^{2r}}{\Delta - \max(z) - \rho}\left(\frac{\rho}{\Delta - \max(z)}\right)^2 \\
&= \Theta(\Delta^{2r-1+2\max\{1-2r,2r-2\}}) = \Theta(\Delta^{\max\{1-2r,6r-5\}})
\end{aligned} \tag{2.49}
$$

where $\rho = \|\mathbf{H}_{\text{else}}\| + 2^{-1/3}\alpha^{1/3}\Delta^{2-2r} + 2^{1/3}\alpha^{2/3}\Delta^{2r-1}$. If we again write the self energy expansion Equation 2.48 as

$$\mathbf{\Sigma}_-(z) = \mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w + \Theta(\Delta^{f(r)}),$$

the function $f(r) < 0$, which determines the dominant power in Δ among $\mathbf{E}_1$ through $\mathbf{E}_6$, can be found as

$$f(r) = \max\{1-2r, 2r-2, 4r-3, 6r-5\}, \quad \frac{1}{2} < r < 1. \tag{2.50}$$

Similar to the discussion after Equation 2.40, the optimal scaling of $\Delta = \Theta(\epsilon^{1/f(r)})$ gives $r = \text{argmin} f(r) = 2/3$, when $f(r) = -1/3$ and $\Delta = \Theta(\epsilon^{-3})$, as is shown in Figure 2.5a. Note that the $4r-3$ component in $f(r)$, Equation 2.50, comes from the error term $\mathbf{E}_4$ in Equation 2.48. The idea for improving the gadget construction comes from the observation in Figure 2.5a that when we add a term in $\mathbf{V}$ to compensate for $\mathbf{E}_4$, the dominant power of Δ in the perturbation series, $f(r)$, could admit a lower minimum as shown in Figure 2.5b. In the previous calculation we have shown that this is indeed the case and the minimum value of $f(r)$ becomes $-1/2$ in the improved case, indicating that $\Delta = \Theta(\epsilon^{-2})$ is sufficient for keeping the error terms $O(\epsilon)$.

2.5 Parallel 3- to 2-body gadget

Summary. In Section 2.3 we have shown that by using parallel subdivision gadgets iteratively, one can reduce a k-body target term to 3-body. We now turn our attention to considering $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \sum_{i=1}^{m} \alpha_i \mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i$, which is a sum of m 3-body terms. A straightforward approach to the reduction is to deal with the 3-body terms in series *i.e.* one at a time: apply a 3-body gadget on one term, and include the entire gadget in the $\mathbf{H}_{\text{else}}$ of the target Hamiltonian in reducing the next 3-body term. In this construction, Δ scales exponentially as a function of m. In order to avoid that overhead, we apply all gadgets in parallel, which means introducing m ancilla spins, one for each 3-body term and applying the same Δ onto it. This poses additional challenges as the operator valued resolvent $\mathbf{G}(z)$ now has multiple poles. Enumerating high order terms in the perturbation series requires consideration of the combinatorial properties of the bit flipping processes (Figure 2.6).

If we apply the current construction [8, 129] of 3-body gadgets in parallel, which requires $\Delta = \Theta(\epsilon^{-3})$, it can be shown [129] that the cross-gadget contribution is

$O(\epsilon)$. However, if we apply our improved construction of the 3- to 2-body gadget in parallel, the perturbation expansion will contain $\Theta(1)$ cross-gadget terms that are dependent on the commutation relations between $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{A}_j$, $\mathbf{B}_j$. Compensation terms are designed to ensure that these error terms are suppressed in the perturbative expansion. With our improved parallel 3-body construction, $\Delta = \Theta(\epsilon^{-2}\text{poly}(m))$ is sufficient.

The combination of parallel subdivision with the parallel 3- to 2-body reduction allows us to reduce an arbitrary k-body target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{S}_1 \mathbf{S}_2 \cdots \mathbf{S}_k$ to 2-body [129]. In this chapter we have improved both parallel 2-body and 3- to 2-body gadgets. When numerically optimized at each iteration, our construction requires a smaller gap than the original construction [129] for the range of k concerned.

Analysis. In Section 2.3 we have shown that with subdivision gadgets one can reduce a k-body interaction term down to 3-body. To complete the discussion on reducing a k-body term to 2-body, now we deal with reducing a 3-body target Hamiltonian of form

$$\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \sum_{i=1}^{m} \alpha_i \mathbf{A}_i \otimes \mathbf{B}_i \otimes \mathbf{C}_i$$

where $\mathbf{H}_{\text{else}}$ is a finite-norm Hamiltonian and all of $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{C}_i$ are single-qubit Pauli operators acting on one of the n qubits that $\mathbf{H}_{\text{targ}}$ acts on. Here without loss of generality, we assume $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{C}_i$ are single-qubit Pauli operators as our construction depends on the commutation relationships among these operators. The Pauli operator assumption ensures that the commutative relationship can be determined efficiently a priori.

We label the n qubits by integers from 1 to n. We assume that in each 3-body term of the target Hamiltonian, $\mathbf{A}_i$, $\mathbf{B}_i$ and $\mathbf{C}_i$ act on three different qubits whose labels are in increasing order i.e. if we label the qubits with integers from 1 to n, $\mathbf{A}_i$ acts on qubit a_i, $\mathbf{B}_i$ acts on b_i, $\mathbf{C}_i$ on c_i, we assume that $1 \leq a_i < b_i < c_i \leq n$ must hold for all values of i from 1 to m.

One important feature of this gadget is that the gap Δ scales as $\Theta(\epsilon^{-2})$ instead of the common $\Theta(\epsilon^{-3})$ scaling assumed by the other 3-body constructions in the literature [7, 8, 129].

To reduce the $\mathbf{H}_{\text{targ}}$ to 2-body, introduce m qubits labelled as u_1, u_2, $\cdots$, u_m and apply an energy penalty Δ onto the excited subspace of each qubit, as in the case of parallel subdivision gadgets presented previously. Then we have

$$\mathbf{H} = \sum_{i=1}^{m} \Delta |1\rangle\langle 1|_{u_i} = \sum_{x \in \{0,1\}^m} h(x) \Delta |x\rangle\langle x|. \tag{2.51}$$

where $h(x)$ is the Hamming weight of the m-bit string x. In this new construction the perturbation V is defined as

$$\mathbf{V} \;=\; \mathbf{H}_{\text{else}} + \sum_{i=1}^{m} \mu_i \mathbf{C}_i \otimes |1\rangle\langle 1|_{u_i} + \sum_{i=1}^{m} (\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i) \otimes \mathbf{X}_{u_i} + \mathbf{V}_1 + \mathbf{V}_2 + \mathbf{V}_3 \tag{2.52}$$

where $\mathbf{V}_1$ is defined as

$$\mathbf{V}_1 = \frac{1}{\Delta} \sum_{i=1}^{m} (\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i)^2 - \frac{1}{\Delta^2} \sum_{i=1}^{m} (\kappa_i^2 + \lambda_i^2) \mu_i \mathbf{C}_i \tag{2.53}$$

and $\mathbf{V}_2$ is defined as

$$\mathbf{V}_2 = -\frac{1}{\Delta^3} \sum_{i=1}^{m} (\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i)^4. \tag{2.54}$$

$\mathbf{V}_3$ will be explained later. Following the discussion in Section 2.4, the coefficients κ_i, λ_i and μ_i are defined as

$$\kappa_i = \text{sgn}(\alpha_i) \left(\frac{|\alpha_i|}{2} \right)^{\frac{1}{3}} \Delta^{\frac{3}{4}}, \quad \lambda_i = \left(\frac{|\alpha_i|}{2} \right)^{\frac{1}{3}} \Delta^{\frac{3}{4}}, \quad \mu_i = \left(\frac{|\alpha_i|}{2} \right)^{\frac{1}{3}} \Delta^{\frac{1}{2}}. \tag{2.55}$$

However, as we will show in detail later in this section, a close examination of the perturbation expansion based on the $\mathbf{V}$ in Equation 2.52 shows that with assignments of κ_i, λ_i and μ_i in Equation 2.55 if $\mathbf{V}$ has only $\mathbf{V}_1$ and $\mathbf{V}_2$ as compensation terms, the cross-gadget contribution in the expansion causes $\Theta(1)$ error terms to arise. In order to compensate for the $\Theta(1)$ error terms, we introduce the compensation

$$\mathbf{V}_3 = \sum_{i=1}^{m} \sum_{j=1, j \neq i}^{m} \bar{\mathbf{V}}_{ij}$$

into $\mathbf{V}$ and $\bar{\mathbf{V}}_{ij}$ is the compensation term for cross-gadget contribution[1]. Before presenting the detailed form of $\bar{\mathbf{V}}_{ij}$, let $s_1^{(i,j)} = s_{11}^{(i,j)} + s_{12}^{(i,j)}$ where

$$s_{11}^{(i,j)} = \begin{cases} 1 & \text{if } \begin{cases} [\mathbf{A}_i, \mathbf{A}_j] \neq 0 \\ [\mathbf{B}_i, \mathbf{B}_j] = 0 \end{cases} \text{ or } \begin{cases} [\mathbf{B}_i, \mathbf{B}_j] \neq 0 \\ [\mathbf{A}_i, \mathbf{A}_j] = 0 \end{cases} \\ 0 & \text{otherwise} \end{cases} \tag{2.56}$$

$$s_{12}^{(i,j)} = \begin{cases} 1 & \text{if } [\mathbf{A}_i, \mathbf{B}_j] \neq 0 \text{ or } [\mathbf{B}_i, \mathbf{A}_j] \neq 0 \\ 0 & \text{otherwise} \end{cases} \tag{2.57}$$

and further define $s_2^{(i,j)}$ as

$$s_2^{(i,j)} = \begin{cases} 1 & \text{if } [\mathbf{A}_i, \mathbf{A}_j] \neq 0 \text{ and } [\mathbf{B}_i, \mathbf{B}_j] \neq 0 \\ 0 & \text{otherwise.} \end{cases} \tag{2.58}$$

Then we define $\bar{\mathbf{V}}_{ij}$ as

$$\bar{\mathbf{V}}_{ij} \;=\; -s_1^{(i,j)} \cdot \frac{1}{\Delta^3}(\kappa_i\kappa_j)^2\mathbf{I} - s_2^{(i,j)}\left(\frac{2}{\Delta^3}(\kappa_i\kappa_j)^2\mathbf{I} - \frac{2}{\Delta^3}\kappa_i\kappa_j\lambda_i\lambda_j\mathbf{A}_i\mathbf{A}_j\mathbf{B}_i\mathbf{B}_j\right) \tag{2.59}$$

where $s_1^{(i,j)}$ and $s_2^{(i,j)}$ are coefficients that depend on the commuting relations between the operators in the i^{th} term and the j^{th} term. Note that in Equation 2.59, although the term $\mathbf{A}_i\mathbf{A}_j\mathbf{B}_i\mathbf{B}_j$ is 4-local, it arises only in cases where $s_2^{(i,j)} = 1$. In this case, an additional gadget with a new ancilla u_{ij} can be introduced to generate the 4-local term. For succinctness we present the details of this construction in Appendix A. With the penalty Hamiltonian $\mathbf{H}$ defined in Equation 2.51, the operator-valued resolvent (or the Green's function) can be written as

$$\mathbf{G}(z) = \sum_{x\in\{0,1\}^m} \frac{1}{z - h(x)\Delta}|x\rangle\langle x|. \tag{2.60}$$

[1]As is shown by [129], for the gadget construction with the assignments of κ_i, λ_i and μ_i all being $O(\Delta^{2/3})$, the cross-gadget contribution can be reduced by increasing Δ, thus no cross-gadget compensation is needed. However, with our assignments of κ_i, λ_i and μ_i in (2.55) there are cross-gadget error terms in the perturbative expansion that are of order $O(1)$, which cannot be reduced by increasing Δ. This is why we need $\bar{V}_{ij}$. Since the $O(1)$ error terms are dependent on the commuting relations between A_i, B_i, A_j and B_j of each pair of i^{th} and j^{th} terms in the target Hamiltonian, $\bar{V}_{ij}$ depends on their commutation relations too.

Define subspaces of the ancilla register $\mathcal{L}_- = \text{span}\{|00\cdots0\rangle\}$ and $\mathcal{L}_+ = \text{span}\{|x\rangle | x \neq 00\cdots0\}$. Define $\mathbf{P}_-$ and $\mathbf{P}_+$ as the projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$. Then the projections of $\mathbf{V}$ onto the subspaces can be written as

$$\begin{aligned}
\mathbf{V}_+ &= \Bigg(H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)^2 - \frac{1}{\Delta^2}\sum_{i=1}^{m}(\kappa_i^2+\lambda_i^2)\mu_i\mathbf{C}_i \\
&- \frac{1}{\Delta^3}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)^4 + \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m}\bar{\mathbf{V}}_{ij}\Bigg)\otimes\mathbf{P}_+ + \sum_{i=1}^{m}\mu_i C_i\otimes\mathbf{P}_+|1\rangle\langle 1|_{u_i}\mathbf{P}_+ \\
&+ \underbrace{\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)\otimes\mathbf{P}_+\mathbf{X}_{u_i}\mathbf{P}_+}_{\mathbf{V}_f} = \mathbf{V}_s + \mathbf{V}_f \\
\mathbf{V}_{-+} &= \sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)\otimes\mathbf{P}_-\mathbf{X}_{u_i}\mathbf{P}_+, \quad \mathbf{V}_{+-} = \sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)\otimes\mathbf{P}_+\mathbf{X}_{u_i}\mathbf{P}_- \\
\mathbf{V}_- &= \Bigg(\mathbf{H}_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)^2 - \frac{1}{\Delta^2}\sum_{i=1}^{m}(\kappa_i^2+\lambda_i^2)\mu_i\mathbf{C}_i \\
&- \frac{1}{\Delta^3}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)^4 + \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m}\bar{\mathbf{V}}_{ij}\Bigg)\otimes\mathbf{P}_-.
\end{aligned} \tag{2.61}$$

Here the $\mathbf{V}_+$ projection is intentionally divided up into $\mathbf{V}_f$ and $\mathbf{V}_s$ components. $\mathbf{V}_f$ is the component of $\mathbf{V}_+$ that contributes to the perturbative expansion only when the perturbative term corresponds to flipping processes in the $\mathcal{L}_+$ subspace. $\mathbf{V}_s$ is the component that contributes only when the perturbative term corresponds to transitions that involve the state of the m-qubit ancilla register staying the same.

The projection of the Green's function $\mathbf{G}(z)$ onto $\mathcal{L}_+$ can be written as

$$\mathbf{G}_+(z) = \sum_{x\neq 0\cdots00}\frac{1}{z - h(x)\Delta}|x\rangle\langle x|. \tag{2.62}$$

We now explain the self energy expansion

$$\begin{aligned}
\mathbf{\Sigma}_-(z) &= \mathbf{V}_- + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-} + \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^2\mathbf{G}_+\mathbf{V}_{+-} \\
&+ \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^3\mathbf{G}_+\mathbf{V}_{+-} + \cdots
\end{aligned} \tag{2.63}$$

in detail term by term. The 1st order term is simply $\mathbf{V}_-$ from Equation Equation 2.61. The 2nd order term corresponds to processes of starting from an all-zero state of the m ancilla qubits, flipping one qubit and then flipping it back:

$$\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} = \frac{1}{z-\Delta}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)^2 \tag{2.64}$$

The 3rd order term corresponds to processes of starting from an all-zero state of the ancilla register, flipping one qubit, staying at the same state for $\mathbf{V}_+$ and then flipping the same qubit back. Therefore only the $\mathbf{V}_f$ component in $\mathbf{V}_+$ in Equation Equation 2.61 will contribute to the perturbative expansion for $\mathbf{T}_3 = \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-}$:

$$\begin{aligned}\mathbf{T}_3 &= \frac{1}{(z-\Delta)^2}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i)\Bigg[H_{\text{else}} + \mu_i\mathbf{C}_i + \frac{1}{\Delta}\sum_{j=1}^{m}(\kappa_j\mathbf{A}_j + \lambda_j\mathbf{B}_j)^2 \\ &+ \frac{1}{\Delta^2}\sum_{j=1}^{m}\Bigg[(\kappa_j^2+\lambda_j^2)\mu_j\mathbf{C}_j - \frac{1}{\Delta^3}\sum_{j=1}^{m}(\kappa_j\mathbf{A}_j + \lambda_j\mathbf{B}_j)^4 + \sum_{j=1}^{m}\sum_{l=1,l\neq j}^{m}\bar{\mathbf{V}}_{jl}\Bigg] \\ &\quad (\kappa_i\mathbf{A}_i + \lambda_i\mathbf{B}_i).\end{aligned} \tag{2.65}$$

The 4th order term is more involved. Here we consider two types of transition processes (for diagrammatic illustration refer to Figure 2.6):

1. Starting from the all-zero state, flipping one of the qubits, flipping another qubit, then using the remaining $\mathbf{V}_+$ and $\mathbf{V}_{+-}$ to flip both qubits back one after the other (there are 2 different possible sequences, see Figure 2.6a).

2. Starting from the all-zero state of the ancilla register, flipping one of the qubits, staying twice for the two $\mathbf{V}_+$ components and finally flipping back the qubit during $\mathbf{V}_{+-}$ (Figure 2.6b).

Therefore in the transition processes of type (1), $\mathbf{V}_+$ will only contribute its $\mathbf{V}_f$ component and the detailed form of its contribution depends on which qubit in the ancilla register is flipped. The two possibilities of flipping the two qubits back explains why the second term in Equation 2.66 takes the form of a summation of two

components. Because two qubits are flipped during the transition, $\mathbf{G}_+$ will contribute a $\frac{1}{z-2\Delta}$ factor and two $\frac{1}{z-\Delta}$ factors to the perturbative term.

In the transition processes of type (2), $\mathbf{V}_+$ will only contribute its $\mathbf{V}_s$ component to the 4^{th} order term since the states stay the same during both $\mathbf{V}_+$ operators in the perturbative term. $\mathbf{G}_+$ will only contribute a factor of $\frac{1}{z-\Delta}$ because the Hamming weight of the bit string represented by the state of the ancilla register is always 1. This explains the form of the first term $\mathbf{T}_4 = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^2\mathbf{G}_+\mathbf{V}_{+-}$ in Equation 2.66.

$$
\begin{aligned}
\mathbf{T}_4 &= \frac{1}{(z-\Delta)^3}\sum_{i=1}^{m}(\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i)\Bigg[\mathbf{H}_{\text{else}}+\mu_i\mathbf{C}_i+\frac{1}{\Delta}\sum_{j=1}^{m}(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j)^2 \\
&- \frac{1}{\Delta^2}\sum_{j=1}^{m}(\kappa_j^2+\lambda_j^2)\mu_j\mathbf{C}_j-\frac{1}{\Delta^3}\sum_{j=1}^{m}(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j)^4+\sum_{j=1}^{m}\sum_{l=1,l\neq j}^{m}\bar{\mathbf{V}}_{jl}\Bigg]^2 \\
&\quad (\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i) \\
&+ \frac{1}{(z-\Delta)^2(z-2\Delta)}\sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m}\Big[(\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i)(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j) \\
&\qquad\qquad (\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i)(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j) \\
&+ (\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i)(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j)(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j)(\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i)\Big].
\end{aligned}
\tag{2.66}
$$

Although the 4^{th} order does not contain terms that are useful for simulating the 3-body target Hamiltonian, our assignments of κ_i, λ_i and μ_i values in Equation 2.55 imply that some of the terms at this order can be $\Theta(1)$. Indeed, the entire second term in Equation 2.66 is of order $\Theta(1)$ based on Equation 2.55. Therefore it is necessary to study in detail what error terms arise at this order and how to compensate for them in the perturbation $\mathbf{V}$. A detailed analysis on how to compensate the $\Theta(1)$ errors is presented in the Appendix A. In Figure 2.6 we illustrate the transitions that occur at 4th order. Each diagram deals with a fixed pair of ancilla qubits labelled i and j. The diagram (a) has three horizontal layers connected with vertically going arrows. $\mathbf{V}_f$ and $\mathbf{V}_s$ are both components of $\mathbf{V}_+$. In fact $\mathbf{V}_+ = \mathbf{V}_f + \mathbf{V}_s$ where $\mathbf{V}_f$ is responsible for the flipping and $\mathbf{V}_s$ contributes when the transition does not have flipping. At the left of each horizontal layer lies the expression for $\mathbf{G}_+(z)$, which is different for states

in $\mathcal{L}_+$ with different Hamming weights. The diagram (b) is constructed in a similar fashion except that we are dealing with the type of 4th order transition where the state stays the same for two transitions in $\mathcal{L}_+$, hence the $\mathbf{V}_s$ symbols and the arrows going from one state to itself. The diagram (a) reflects the type of 4th order transition that induces cross-gadget contribution and given our gadget parameter setting, this contribution could be $O(1)$ when otherwise compensated. The diagram (b) shows two paths that don not interfere with each other and thus having no cross-gadget contributions. The 5^{th} order and higher terms are errors that can be reduced by increasing Δ:

$$\sum_{k=3}^{\infty} \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^k\mathbf{G}_+\mathbf{V}_{+-}. \tag{2.67}$$

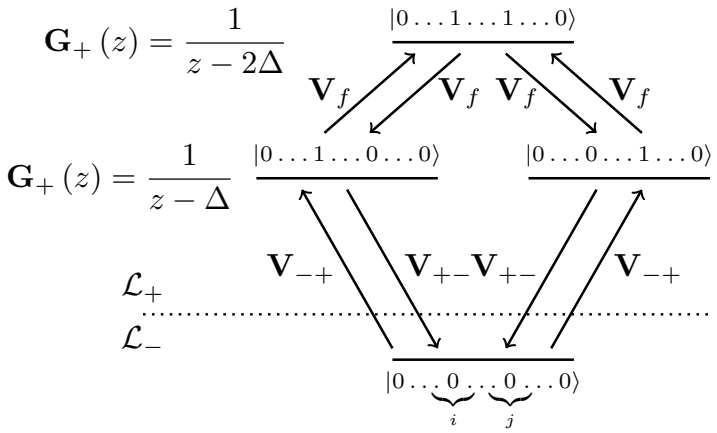
At first glance, with assignments of κ_i, λ_i and μ_i in Equation 2.55, it would appear that this error term is $\Theta(\Delta^{-1/4})$ since $\|\mathbf{V}_{-+}\| = \Theta(\Delta^{3/4})$, $\|\mathbf{V}_{+-}\| = \Theta(\Delta^{3/4})$, $\|\mathbf{V}_+\| = \Theta(\Delta^{3/4})$ and $\|\mathbf{G}_+\| = \Theta(\Delta^{-1})$,

$$\begin{aligned}
\sum_{k=3}^{\infty} \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^k\mathbf{G}_+\mathbf{V}_{+-} &\leq \sum_{k=3}^{\infty} \|\mathbf{V}_{-+}\| \cdot \|\mathbf{G}_+\mathbf{V}_+\|^k \|\mathbf{G}_+\| \cdot \|\mathbf{V}_{+-}\| \\
&= \|\mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^3\mathbf{G}_+\mathbf{V}_{+-}\| \sum_{k=0}^{\infty} \|\mathbf{G}_+\mathbf{V}_+\|^k \\
&= O(\Delta^{-1/4})
\end{aligned} \tag{2.68}$$

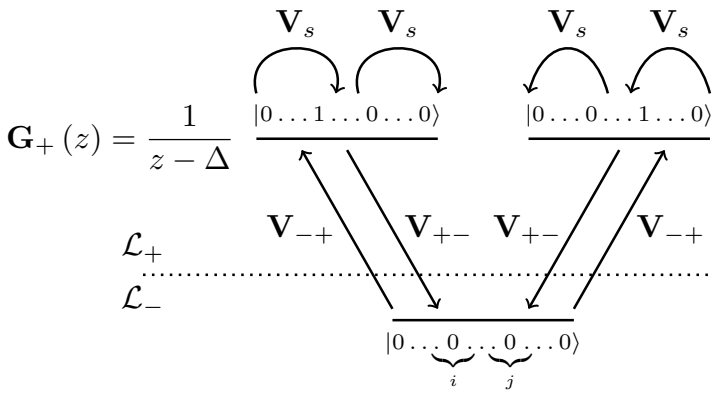
as $\sum_{k=0}^{\infty} \|\mathbf{G}_+\mathbf{V}_+\|^k = O(1)$. However, here we show that in fact this term in Equation 2.67 is $\Theta(\Delta^{-1/2})$. Note that the entire term Equation 2.67 consists of contributions from the transition processes where one starts with a transition from the all-zero state to a state $|x\rangle$ with $x \in \{0,1\}^m$ and $h(x) = 1$. If we focus on the perturbative term of order $k+2$:

$$\mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^k\mathbf{G}_+\mathbf{V}_{+-},$$

after k steps. During every step one can choose to either flip one of the ancilla qubits or stay in the same state of the ancilla register, the state of the ancilla register will go back to a state $|y\rangle$ with $y \in \{0,1\}^m$ and $h(y) = 1$. Finally the $|1\rangle$ qubit in $|y\rangle$ is flipped back to $|0\rangle$ and we are back to the all-zero state which spans the ground state



(a)



(b)

Fig. 2.6. Diagrams illustrating the transitions that occur at 4^{th} order. The two diagrams each represent a type of transition that occurs at 4^{th} order. Each diagram is divided by a horizontal line where below the line is $\mathcal{L}_-$ space and above is $\mathcal{L}_+$ subspace.

subspace $\mathcal{L}_-$. Define the total number of flipping steps to be k_f. Then for a given k, k_f takes only values from

$$K(k) = \begin{cases} \{k, k-2, \cdots, 2\} & \text{if } k \text{ is even} \\ \{k-1, k-3, \cdots, 2\} & \text{if } k \text{ is odd.} \end{cases} \tag{2.69}$$

2.6 Creating 3-body gadget from local X terms

Summary. In general, terms in perturbative gadgets involve mixed couplings (e.g. X_iZ_j). Although such couplings can be realized by certain gadget constructions [33], physical couplings of this type are difficult to realize in an experimental setting. However, there has been significant progress towards experimentally implementing Ising models with transverse fields of the type [135]:

$$\mathbf{H}_{ZZ} = \sum_i \delta_i \mathbf{X}_i + \sum_i h_i \mathbf{Z}_i + \sum_{i,j} J_{ij} \mathbf{Z}_i \mathbf{Z}_j. \tag{2.70}$$

Accordingly, an interesting question is whether we can approximate 3-body terms such as $\alpha \cdot \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}_k$ using a Hamiltonian of this form. This turns out to be possible by employing a perturbative calculation which considers terms up to 5^{th} order.

Similar to the 3- to 2-body reduction discussed previously, we introduce an ancilla w and apply the Hamiltonian $\mathbf{H} = \Delta|1\rangle\langle 1|_w$. We apply the perturbation

$$\mathbf{V} = \mathbf{H}_{\text{else}} + \mu(\mathbf{Z}_i + \mathbf{Z}_j + \mathbf{Z}_k) \otimes |1\rangle\langle 1|_w + \mu \mathbf{I} \otimes \mathbf{X}_w + \mathbf{V}_{\text{comp}} \tag{2.71}$$

where $\mu = (\alpha\Delta^4/6)^{1/5}$ and $\mathbf{V}_{\text{comp}}$ is

$$\begin{aligned} \mathbf{V}_{\text{comp}} &= \frac{\mu^2}{\Delta}|0\rangle\langle 0|_w - \left(\frac{\mu^3}{\Delta^2} + 7\frac{\mu^5}{\Delta^4}\right)(\mathbf{Z}_i + \mathbf{Z}_j + \mathbf{Z}_k) \otimes |0\rangle\langle 0|_w \\ &+ \frac{\mu^4}{\Delta^3}\left(3\mathbf{I} + 2\mathbf{Z}_i\mathbf{Z}_j + 2\mathbf{Z}_i\mathbf{Z}_k + 2\mathbf{Z}_j\mathbf{Z}_k\right). \end{aligned} \tag{2.72}$$

To illustrate the basic idea of the 5^{th} order gadget, define subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$ in the usual way and define $\mathbf{P}_-$ and $\mathbf{P}_+$ as projectors into these respective subspaces. Then the second term in Equation 2.71 with $\otimes|1\rangle\langle 1|_w$ contributes a linear combination $\mu\mathbf{Z}_i + \mu\mathbf{Z}_j + \mu\mathbf{Z}_k$ to $\mathbf{V}_+ = \mathbf{P}_+\mathbf{V}\mathbf{P}_+$. The third term in Equation 2.71 induces a transition between $\mathcal{L}_-$ and $\mathcal{L}_+$ yet since it operates trivially on qubits 1-3, it only contributes a constant μ to the projections $\mathbf{V}_{-+} = \mathbf{P}_-\mathbf{V}\mathbf{P}_+$ and $\mathbf{V}_{+-} = \mathbf{P}_+\mathbf{V}\mathbf{P}_-$. In the perturbative expansion, the 5^{th} order contains a term

$$\frac{\mathbf{V}_{-+}\mathbf{V}_+\mathbf{V}_+\mathbf{V}_+\mathbf{V}_{+-}}{(z-\Delta)^4} = \frac{\mu^5(\mathbf{Z}_i + \mathbf{Z}_j + \mathbf{Z}_k)^3}{(z-\Delta)^4} \tag{2.73}$$

due to the combined the contribution of the second and third term in Equation 2.71. This yields a term proportional to $\alpha \cdot \mathbf{Z}_i \otimes \mathbf{Z}_j \otimes \mathbf{Z}_k$ along with some 2-local error terms. These error terms, combined with the unwanted terms that arise at 1^{st} through 4^{th} order perturbation, are compensated by V_{comp}. Note that terms at 6^{th} order and higher are $\Theta(\Delta^{-1/5})$. This means in order to satisfy the gadget theorem of Kempe *et al.* ([7, Theorem 3], or Theorem I.1) Δ needs to be $\Theta(\epsilon^{-5})$. This is the first perturbative gadget that simulates a 3-body target Hamiltonian using the Hamiltonian Equation 2.70. By rotating the ancilla space, subdivision gadgets can also be implemented using this Hamiltonian: in the $\mathbf{X}$ basis, $\mathbf{Z}$ terms will induce a transition between the two energy levels of $\mathbf{X}$. Therefore $\mathbf{Z}_i\mathbf{Z}_j$ coupling could be used for a perturbation of the form in Equation 2.1 in the rotated basis. In principle using the transverse Ising model in Equation 2.70, one can reduce some diagonal k-body Hamiltonian to 3-body by iteratively applying the subdivision gadget and then to 2-body by using the 3-body reduction gadget.

Analysis. Similar to the gadgets we have presented so far, we introduce an ancilla spin w. Applying an energy gap Δ on the ancilla spin gives the unperturbed Hamiltonian $\mathbf{H} = \Delta|1\rangle\langle 1|_w$. We then perturb the Hamiltonian $\mathbf{H}$ using a perturbation $\mathbf{V}$ described in (2.71). Using the same definitions of subspaces $\mathcal{L}_+$ and $\mathcal{L}_-$ as the previous 3-body gadget, the projections of $\mathbf{V}$ into these subspaces can be written as

$$\mathbf{V}_+ = \left\{ \mathbf{H}_{\text{else}} + \mu(\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3) + \frac{\mu^4}{\Delta^3}\big[3\mathbf{I} + 2(\mathbf{Z}_1\mathbf{Z}_2 + \mathbf{Z}_1\mathbf{Z}_3 + \mathbf{Z}_2\mathbf{Z}_3)\big] \right\} \otimes |1\rangle\langle 1|_w \tag{2.74}$$

$$\mathbf{V}_- = \left\{ \mathbf{H}_{\text{else}} + \frac{\mu^2}{\Delta}\mathbf{I} - \frac{\mu^3}{\Delta^2}(\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3)\mathbf{I} + \frac{\mu^4}{\Delta^3}\big[3\mathbf{I} + 2(\mathbf{Z}_1\mathbf{Z}_2 + \mathbf{Z}_1\mathbf{Z}_3 + \mathbf{Z}_2\mathbf{Z}_3)\big] \right. \tag{2.75}$$

$$\mathbf{V}_{-+} = \mu\mathbf{I} \otimes |0\rangle\langle 1|_w, \quad \mathbf{V}_{+-} = \mu\mathbf{I} \otimes |1\rangle\langle 0|_w. \tag{2.76}$$

The low-lying spectrum of $\widetilde{\mathbf{H}}$ is approximated by the self energy expansion $\mathbf{\Sigma}_-(z)$ below with $z \in [-\max z, \max z]$ where $\max z = \|\mathbf{H}_{\text{else}}\| + |\alpha| + \epsilon$. With the choice of μ above the expression of $\mathbf{V}_+$ in Equation 2.76 can be written as

$$\mathbf{V}_+ = \left(\mathbf{H}_{\text{else}} + \mu(\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3) + O(\Delta^{1/5})\right) \otimes |1\rangle\langle 1|_w. \tag{2.77}$$

Because we are looking for the 5^{th} order term in the perturbation expansion that gives a term proportional to $\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_3$, expand the self energy in Equation 1.26 up to 5^{th} order:

$$\begin{aligned}
\mathbf{\Sigma}_-(z) &= \mathbf{V}_- \otimes |0\rangle\langle 0|_w + \frac{\mathbf{V}_{-+}\mathbf{V}_{+-}}{z-\Delta} \otimes |0\rangle\langle 0|_w + \frac{\mathbf{V}_{-+}\mathbf{V}_{+}\mathbf{V}_{+-}}{(z-\Delta)^2} \otimes |0\rangle\langle 0|_w \\
&+ + \frac{\mathbf{V}_{-+}\mathbf{V}_{+}\mathbf{V}_{+}\mathbf{V}_{+-}}{(z-\Delta)^3} \otimes |0\rangle\langle 0|_w + \frac{\mathbf{V}_{-+}\mathbf{V}_{+}\mathbf{V}_{+}\mathbf{V}_{+}\mathbf{V}_{+-}}{(z-\Delta)^4} \otimes |0\rangle\langle 0|_w \\
&+ \sum_{k=4}^{\infty} \frac{\mathbf{V}_{-+}\mathbf{V}_{+}^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}} \otimes |0\rangle\langle 0|_w.
\end{aligned} \tag{2.78}$$

Using this simplification as well as the expressions for $\mathbf{V}_-$, $\mathbf{V}_{-+}$ and $\mathbf{V}_{+-}$ in Equation 2.76, the self energy expansion Equation 2.78 up to 5^{th} order becomes

$$\begin{aligned}
\mathbf{\Sigma}_-(z) = &\underbrace{\left(\mathbf{H}_{\text{else}} + \frac{6\mu^5}{\Delta^4}\mathbf{Z}_1\mathbf{Z}_2\mathbf{Z}_3\right) \otimes |0\rangle\langle 0|_w}_{\mathbf{H}_{\text{eff}}} + \underbrace{\left(\frac{1}{\Delta} + \frac{1}{z-\Delta}\right)\mu^2\mathbf{I} \otimes |0\rangle\langle 0|_w}_{\mathbf{E}_1} \\
&+ \underbrace{\left(\frac{1}{(z-\Delta)^2} - \frac{1}{\Delta^2}\right)\mu^3(\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3) \otimes |0\rangle\langle 0|_w}_{\mathbf{E}_2} \\
&+ \underbrace{\left(\frac{1}{\Delta^3} + \frac{1}{(z-\Delta)^3}\right) \cdot \mu^4 \cdot (\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3)^2 \otimes |0\rangle\langle 0|_w}_{\mathbf{E}_3} \\
&+ \underbrace{\left(\frac{1}{(z-\Delta)^4} - \frac{1}{\Delta^4}\right)7\mu^5(\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3) \otimes |0\rangle\langle 0|_w}_{\mathbf{E}_4} \\
&+ \underbrace{\frac{\mu^2}{(z-\Delta)^2} \cdot \frac{\mu^4}{\Delta^3}(\mathbf{Z}_1 + \mathbf{Z}_2 + \mathbf{Z}_3)^2 \otimes |0\rangle\langle 0|_w}_{\mathbf{E}_6} \\
&+ O(\Delta^{-2/5}) + O(\|\mathbf{H}_{\text{else}}\|\Delta^{-2/5}) + O(\|\mathbf{H}_{\text{else}}\|^2\Delta^{-7/5}) + O(\|\mathbf{H}_{\text{else}}\|^3\Delta^{-12/5})
\end{aligned}$$

$$+\underbrace{\sum_{k=4}^{\infty}\frac{\mathbf{V}_{-+}\mathbf{V}_{+}^{k}\mathbf{V}_{+-}}{(z-\Delta)^{k+1}}\otimes|0\rangle\langle 0|_{w}}_{\mathbf{E}_7}\,. \tag{2.79}$$

Similar to what we have done in the previous sections, the norm of the error terms $\mathbf{E}_1$ through $\mathbf{E}_7$ can be bounded from above by letting $z \mapsto \max z$. Then we find that

$$\|\mathbf{\Sigma}_{-}(z)-\mathbf{H}_{\text{targ}}\otimes|0\rangle\langle 0|_{w}\| \;\leq\; \Theta(\Delta^{-1/5}) \tag{2.80}$$

if we only consider the dominant dependence on Δ and regard $\|\mathbf{H}_{\text{else}}\|$ as a given constant. To guarantee that $\|\mathbf{\Sigma}_{-}(z)-\mathbf{H}_{\text{targ}}\otimes|0\rangle\langle 0|_{w}\| \leq \epsilon$, we let the right hand side of Equation 2.80 to be $\leq \epsilon$, which translates to $\Delta = \Theta(\epsilon^{-5})$.

This $\Theta(\epsilon^{-5})$ scaling is numerically illustrated (Figure 2.7a). Although in principle the 5^{th} order gadget can be implemented on a Hamiltonian of form Equation 2.70, for a small range of α, the minimum Δ needed is already large (Figure 2.7b), rendering it challenging to demonstrate the gadget experimentally with current resources. However, this is the only currently known gadget realizable with a transverse Ising model that is able to address the case where $\mathbf{H}_{\text{else}}$ is not necessarily diagonal.

2.7 Alternative construction for k- to 2-body reduction

Summary. We have presented in Section 2.3 a general method of reducing k-body interactions to 3-body ones. Subsequently, we used parallel 3-body gadgets in Section 2.5 to reduce k-body interactions to 2-body ones. There is an alternative method, originally due to Kempe, Kitaev and Regev [7] and later generalized by Jordan and Farhi [9] for the k- to 2-body reduction. Unlike the gadgets used in Sections 2.3 and 2.5, which iteratively reduces k-body terms to 2-body and introduces one ancilla per reduction, the gadgets in [7,9] directly reduce a k-body term to 2-body by introducing k ancillas and constructing the unperturbed Hamiltonian $\mathbf{H}$ over the ancilla register — such that its ground state subspace is spanned by all-zero and all-one states. The perturbation $\mathbf{V}$ is constructed such that when the spectrum of $\mathbf{H}$ is perturbed by $\mathbf{V}$, the lowest order non-trivial contribution in the perturbed expansion comes from the

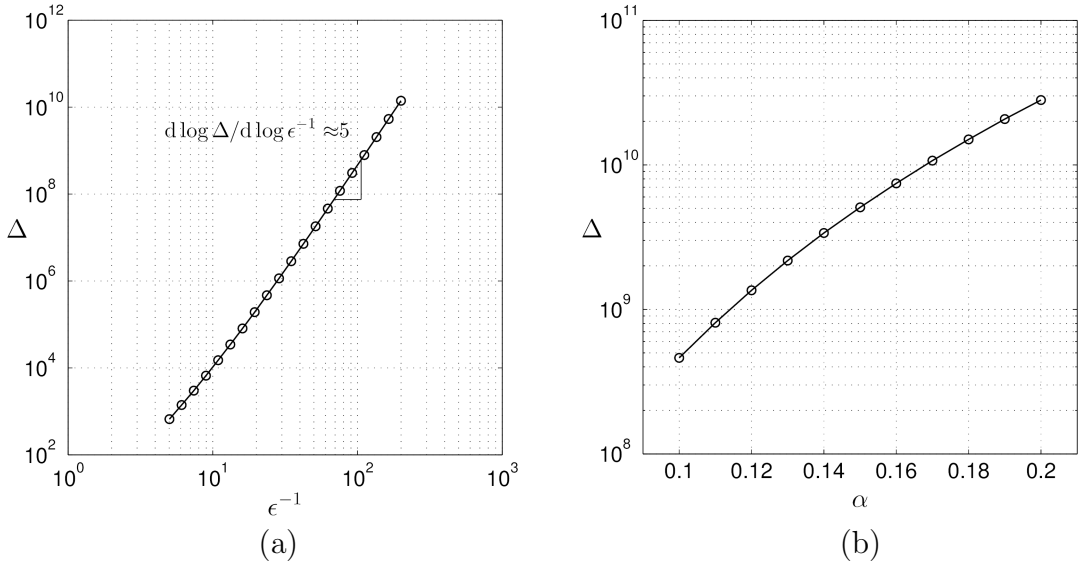


Fig. 2.7. (a): The scaling of minimum Δ needed to ensure $\|\mathbf{\Sigma}_{-}(z) - \mathbf{H}_{\text{eff}}\| \leq \epsilon$ as a function of ϵ^{-1}. Here we choose $\|\mathbf{H}_{\text{else}}\| = 0$, $\alpha = 0.1$ and ϵ ranging from $10^{-0.7}$ to $10^{-2.3}$. The values of minimum Δ are numerically optimized. The notion of "optimized case" refers to the search for the gap Δ needed for yielding a spectral error of precisely ϵ between gadget and target Hamiltonian, which is described in Section 2.2. The slope of the line at large ϵ^{-1} is $4.97 \approx 5$, which provides evidence that with the assignments of $\mu = (\alpha\Delta^4/6)^{1/5}$, the optimal scaling of Δ is $\Theta(\epsilon^{-5})$. (b): The numerically optimized gap versus the desired coupling α in the target Hamiltonian. Here $\epsilon = 0.01$ and $\|\mathbf{H}_{\text{else}}\| = 0$.

transitions in the ancilla space that start from all-zero states, flip all of the qubits and terminate in all-one states (or vice versa, start from all-one and end in all-zero). For a k-body target Hamiltonian, the lowest order non-trivial contribution comes right at the k^{th} order. Any lower order terms either vanish or are proportional to identity.

The k-body version of this construction [9] is analyzed based on a formulation of perturbation theory different from that introduced in Section 1.3 and elsewhere [8, 129]. However, for completeness, we review this construction here, along with some new characterization of the construction, such that the optimal scaling of the gap parameter as a function of k.

Analysis. The parallel subdivision (Section 2.3) and the parallel 3-body (Section 2.5) gadgets that we have presented so far enable us to reduce general k-body target Hamiltonians $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{S}_1 \mathbf{S}_2 \cdots \mathbf{S}_k$ to 2-body with $O(k)$ ancilla qubits. There has also been an alternative construction proposed by Kempe, Kitaev and Regev [7] for 3- to 2-body reduction and generalized to k- to 2-body reduction by Jordan and Farhi [9]. In this section we will review the the construction by Jordan and Farhi, reproducing and elaborating on some of the results in the original paper [9].

For simulating a k-body interaction $\mathbf{H}_{\text{targ}} = \alpha \mathbf{S}_1 \mathbf{S}_2 \cdots \mathbf{S}_k$, the construction in [9] introduces k ancilla qubits (labelled as $w_1 \cdots w_k$) and defines the penalty Hamiltonian $\mathbf{H}$ as a 2-body Hamiltonian

$$\mathbf{H} = \frac{\Delta}{2(k-1)} \sum_{1 \le i < j \le k} (\mathbf{I} - \mathbf{Z}_{w_i} \mathbf{Z}_{w_j})$$

applied onto the ancilla qubits whose ground state subspace of $\mathbf{H}$ is

$$\text{span}\{|00\cdots0\rangle, |11\cdots1\rangle\}$$

and all the other states receive an energy penalty of at least Δ. Define $\mathbf{P}_0 = |00\cdots0\rangle\langle 00\cdots0| + |11\cdots1\rangle\langle 11\cdots1|$ as the projector onto the ground state subspace

of $\mathbf{H}$. Define $\mathbf{P}_j$ as the projector into the subspace of $\mathbf{H}$ with energy $\mathbf{E}_j$. Define the perturbation $\mathbf{V}$ as (here the notation is consistent with Section 1.3.1)

$$\mathbf{V} = \lambda \sum_{i=1}^{k} \mathbf{S}_i \otimes \mathbf{X}_i = \lambda \hat{\mathbf{V}}.$$

The gadget Hamiltonian is then defined as $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V} = \mathbf{H} + \lambda \hat{\mathbf{V}}$ where λ is a small parameter. Define $\mathbf{H}_{\text{eff}}$ as the restriction of the gadget Hamiltonian to the lowest 2^k states and further restricted to the subspace spanned by the $\frac{1}{\sqrt{2}}(|00\cdots0\rangle + |11\cdots1\rangle)$ state of the ancilla register. It turns out that using an expansion due to Bloch [9] one can obtain the expression of the effective Hamiltonian as

$$\mathbf{H}_{\text{eff}} = \mathcal{U}\mathcal{A}\mathcal{U}^{-1} \tag{2.81}$$

where both the operators $\mathcal{U}$ and $\mathcal{A}$ have perturbative expansions

$$\mathcal{U} = \sum_{m=0}^{\infty} \mathcal{U}^{(m)} \qquad\qquad \mathcal{A} = \sum_{m=1}^{\infty} \mathcal{A}^{(m)}. \tag{2.82}$$

Here $\mathcal{U}^{(0)} = \mathbf{P}_0$ and $\mathcal{A}^{(1)} = \lambda \mathbf{P}_0 \hat{\mathbf{V}} \mathbf{P}_0$, $\mathcal{U}^{(m)}$ for $m \geq 1$ and $\mathcal{A}^{(m)}$ for $m \geq 2$ are defined as

$$\begin{aligned} \mathcal{U}^{(m)} &= \lambda^m \sum_{(m)} \mathbf{S}^{l_1} \hat{\mathbf{V}} \mathbf{S}^{l_2} \hat{\mathbf{V}} \cdots \mathbf{S}^{l_m} \hat{\mathbf{V}} \mathbf{P}_0 \\ \mathcal{A}^{(m)} &= \lambda^m \sum_{(m-1)} \mathbf{P}_0 \hat{\mathbf{V}} \mathbf{S}^{l_1} \hat{\mathbf{V}} \mathbf{S}^{l_2} \cdot \hat{\mathbf{V}} \mathbf{S}^{l_{m-1}} \hat{\mathbf{V}} \mathbf{P}_0 \end{aligned} \tag{2.83}$$

where

$$\mathbf{S}^l = \begin{cases} \sum_{j \neq 0} \dfrac{\mathbf{P}_j}{(-E_j)^l} & l > 0 \\ -\mathbf{P}_0 & l = 0 \end{cases}$$

and $\sum_{(m)}$ is a sum over a set of m-tuples $(l_1, l_2, \cdots, l_m)$ such that

$$\begin{aligned} l_1 + l_2 + \cdots + l_m &= m \\ l_1 + \cdots + l_p &\geq p \quad (p = 1, 2, \cdots, m-1). \end{aligned} \tag{2.84}$$

Let $|\psi_1\rangle$, $|\psi_2\rangle$, $\cdots |\psi_d\rangle$ be the lowest $d = 2^{k+1}$ eigenvectors of the gadget Hamiltonian $\widetilde{\mathbf{H}}$. Define the projector

$$\mathbf{\Pi} = \sum_{j=1}^{d} |\psi_j\rangle\langle\psi_j|$$

and $\mathbf{\Pi}$ satisfies

$$\mathbf{\Pi} = \mathcal{U}\mathbf{P}_0\mathcal{U}^{-1}.$$

With the above definitions in place, it can be shown (see [9] for details) that the effective Hamiltonian $\mathbf{H}_{\text{eff}}$ takes the form

$$\begin{aligned}\mathbf{H}_{\text{eff}} &= f(\lambda)\mathbf{\Pi} + \left(\frac{k-1}{\Delta}\right)^{k-1}\frac{-k(-\lambda)^k}{(k-1)!}\mathbf{H}_{\text{targ}} \otimes (|0...0\rangle\langle 1...1| + |1...1\rangle\langle 0...0|) \\ &+ O\left(\frac{\lambda^{k+1}}{\Delta^k}\right)\end{aligned} \tag{2.85}$$

where $f(\lambda)$ is a function that can be evaluated analytically. Therefore the effective coupling of the target Hamiltonian

$$\alpha = \left(\frac{k-1}{\Delta}\right)^{k-1}\frac{-k(-\lambda)^k}{(k-1)!} = O\left(\frac{\lambda^k}{\Delta^{k-1}}\right). \tag{2.86}$$

In the approximation Equation 2.97, we require that the high order error terms be $O(\epsilon)$ while the desired k^{th} order term be $\alpha = O(1)$. Hence we require

$$\begin{aligned}\frac{\lambda^{k+1}}{\Delta^k} &= O(\epsilon) \\ \frac{\lambda^k}{\Delta^{k-1}} &= \Theta(1).\end{aligned} \tag{2.87}$$

From the second constraint we have $\lambda = \Theta(\Delta^{(k-1)/k})$. Substituting into the first constraint gives us

$$\frac{\lambda^{k+1}}{\Delta^k} = \Theta\left(\Delta^{\frac{k-1}{k}\cdot(k+1)}/\Delta^k\right) = \Theta(\Delta^{-1/k}) = O(\epsilon) \tag{2.88}$$

With $\Theta(\Delta^{-1/k}) = O(\epsilon)$ we have $\Delta = \Omega(\epsilon^{-k})$, indicating that in order to simulate $\mathbf{H}_{\text{targ}} = \alpha\mathbf{S}_1 \cdots \mathbf{S}_k$ up to absolute error ϵ, the gap Δ in the penalty Hamiltonian scales exponentially as a function of k.

2.7.1 Numerical examples

We use numerics to verify Equation 2.97 for several sample cases. In order to be consistent with the original work by Jordan and Farhi [9], let $\Delta = k - 1$ for all the cases. Define the effective Hamiltonian with energy shift as

$$\widetilde{\mathbf{H}}_{\text{eff}} = \mathbf{H}_{\text{eff}} - f(\lambda)\mathbf{\Pi}. \tag{2.89}$$

Define the ideal term in the expansion Equation 2.97 as

$$\mathbf{H}_{\text{id}} = \frac{-k(-\lambda)^k}{(k-1)!}\mathbf{H}_{\text{targ}} \otimes (|0...0\rangle\langle 1...1| + |1...1\rangle\langle 0...0|).$$

Example: XYZ gadget. As a first example, consider simulating $\mathbf{H}_{\text{targ}} = \mathbf{X}_1\mathbf{Y}_2\mathbf{Z}_3$. Construct the gadget Hamiltonian as described previously [2] and the expansion of $\mathbf{H}_{\text{eff}}$ takes the form of

$$\mathbf{H}_{\text{eff}} = (\mathbf{P}_0 + O(\lambda))\left(\mathcal{A}^{(\leq 2)} + \mathcal{A}^{(3)} + O(\lambda^4)\right)(\mathbf{P}_0 + O(\lambda)) \tag{2.90}$$

where $\mathcal{A}^{(\leq 2)} = \mathcal{A}^{(1)} + \mathcal{A}^{(2)}$ can be evaluated according to the definition in Equation 2.83 as

$$\mathcal{A}^{(\leq 2)} = \lambda\mathbf{P}_0\hat{\mathbf{V}}\mathbf{P}_0 + \lambda^2\mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{P}_0 = -\lambda^2\mathbf{P}_0\hat{\mathbf{V}}\left(\frac{1}{2}\mathbf{P}_1\right)\hat{\mathbf{V}}\mathbf{P}_0 = -\frac{3}{2}\lambda^2\mathbf{P}_0. \tag{2.91}$$

Here the first term $\mathbf{P}_0\hat{\mathbf{V}}\mathbf{P}_0 = 0$ because $\mathbf{P}_0$ projects into the subspace

$$\mathcal{C}_0 = \text{span}\{|000\rangle, |111\rangle\}$$

and the action of $\hat{\mathbf{V}}$ on any state of $\mathcal{C}_0$ will always produce a new state that is not in $\mathcal{C}_0$. The $\mathcal{A}^{(3)}$ term can then be evaluated as

$$\begin{aligned}\mathcal{A}^{(3)} &= \lambda^3(\mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{P}_0 + \mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^2\hat{\mathbf{V}}\mathbf{P}_0\hat{\mathbf{V}}\mathbf{P}_0) \\ &= = \lambda^3\mathbf{P}_0\hat{\mathbf{V}}\left(\tfrac{1}{-2}\mathbf{P}_1\right)\hat{\mathbf{V}}\left(\tfrac{1}{-2}\mathbf{P}_1\right)\hat{\mathbf{V}}\mathbf{P}_0 = \tfrac{1}{4}\lambda^3\mathbf{P}_0\hat{\mathbf{V}}\mathbf{P}_1\hat{\mathbf{V}}\mathbf{P}_1\hat{\mathbf{V}}\mathbf{P}_0.\end{aligned} \tag{2.92}$$

The second term in the first line vanishes because $\mathbf{P}_0\hat{\mathbf{V}}\mathbf{P}_0 = 0$ as shown before.

[2] In fact the gadget construction by Jordan and Farhi for simulating 3-body interaction is identical to that of Kempe, Kitaev and Regev [7].

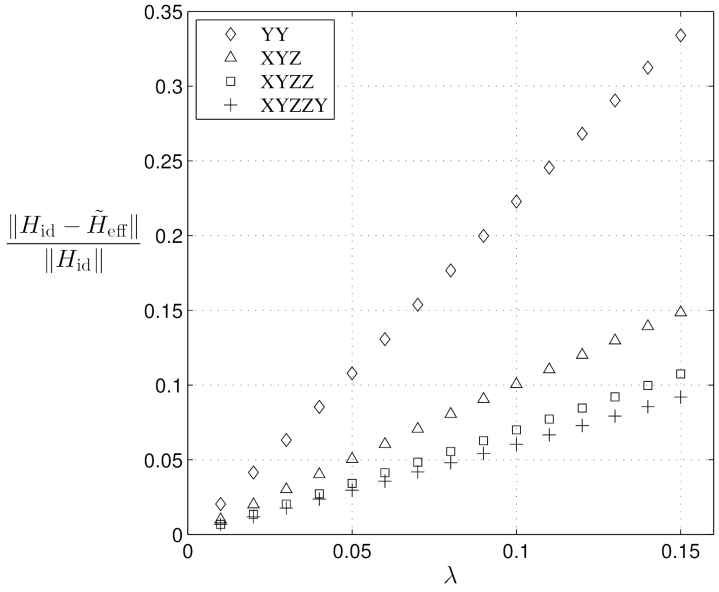


Fig. 2.8. The ratio of the error terms to the ideal Hamiltonian $\mathbf{H}_{\rm id}$ as a function of λ. The XYZ and XYZZ cases are chosen to verify the results by Jordan and Farhi [9]. The YY and XYZZY cases are also plotted.

Define the subspace

$$\mathcal{C}_1 = \text{span}\{|001\rangle, |010\rangle, |100\rangle, |011\rangle, |101\rangle, |110\rangle\}$$

which is the subspace that $\mathbf{P}_1$ projects into. In order for the term $\mathcal{A}^{(3)}$ to be non-zero, the sequential action of the three $\hat{\mathbf{V}}$ terms must first flip one bit to transform a state from $\mathcal{C}_0$ into $\mathcal{C}_1$, then flip one bit and make sure the state still stays in $\mathcal{C}_1$, and finally flip one bit to transform the state back to $\mathcal{C}_0$. For either state in $\mathcal{C}_0$, the only way to accomplish this is to first flip any one of the three bits (there are 3 different ways to do it), then flip a different bit (there are 2 choices) and finally flip the other bit so that all three bits are flipped at the end of the sequence. By elementary combinatorics, there are in total $3 \cdot 2 = 6$ ways to flip all three bits. Therefore the $\mathcal{A}^{(3)}$ term becomes

$$\mathcal{A}^{(3)} = \frac{1}{4}\lambda^3 \cdot 6\mathbf{H}_{\text{targ}} \otimes (|000\rangle\langle 111| + |111\rangle\langle 000|), \tag{2.93}$$

which is our $\mathbf{H}_{\text{id}}$ for the XYZ gadget if we compare it to the Equation 2.97. Then the expansion of $\mathbf{H}_{\text{eff}}$ becomes

$$\mathbf{H}_{\text{eff}} = -\frac{3}{2}\lambda^2\mathbf{\Pi} + \underbrace{\frac{3}{2}\lambda^3\mathbf{H}_{\text{targ}} \otimes (|000\rangle\langle 111| + |111\rangle\langle 000|)}_{\mathbf{H}_{\text{id}}} + O(\lambda^4). \tag{2.94}$$

Hence $f(\lambda) = -\frac{3}{2}\lambda^2$ for the XYZ gadget. Plotting the relative error, which can be expressed as $(\|\widetilde{\mathbf{H}}_{\text{eff}} - \mathbf{H}_{\text{id}}\|)/\|\mathbf{H}_{\text{id}}\|$, versus λ, we have a linear dependence on λ (Figure 2.8). This is expected since the error is $O(\lambda^4)$ and $\mathbf{H}_{\text{id}}$ is $O(\lambda^3)$.

Other examples. Similar analysis can be done for simulating other Hamiltonians from 2-body to 5-body ones. Table 2.1 lists the analytical expressions for all example cases considered and Figure 2.8 plots the relative error for all the cases.

2.7.2 Error analysis

In order to improve the relative error $(\|\widetilde{\mathbf{H}}_{\text{eff}} - \mathbf{H}_{\text{id}}\|)/\|\mathbf{H}_{\text{id}}\|$, it is necessary to examine the expansion terms of higher order in λ than $\mathbf{H}_{\text{id}}$.

Example: 3-body gadget. Consider the $\Theta(\lambda^4)$ terms in the expansion Equation 2.90 for a 3-body gadget simulating $\mathbf{S}_1\mathbf{S}_2\mathbf{S}_3$ where $\mathbf{S}_i$ is any Pauli operator. Expand $\mathcal{U} = \mathbf{P}_0 + \lambda \mathbf{S}^1 \hat{\mathbf{V}} \mathbf{P}_0 + O(\lambda^2)$ up to order λ. Then the expansion Equation 2.90 up to order λ^4 can be written as

$$\begin{aligned}
\mathbf{H}_{\text{eff}} &= -\frac{3}{2}\lambda^2 \mathbf{\Pi} + \underbrace{\frac{3}{2}\lambda^3 \mathbf{H}_{\text{targ}} \otimes (|000\rangle\langle 111| + |111\rangle\langle 000|)}_{\mathbf{H}_{\text{id}}} \\
&- \frac{3}{8}\lambda^4 \mathbf{\Pi} - \frac{3}{4}\lambda^4 \big[\mathbf{S}_2\mathbf{S}_3 \otimes (|00\rangle\langle 11| + |11\rangle\langle 00|)_{w_2 w_3} \\
&\qquad + \mathbf{S}_1\mathbf{S}_3 \otimes (|00\rangle\langle 11| + |11\rangle\langle 00|)_{w_1 w_3} \\
&\qquad \underbrace{+ \mathbf{S}_1\mathbf{S}_2 \otimes (|00\rangle\langle 11| + |11\rangle\langle 00|)_{w_1 w_2}\big]}_{\Theta(\lambda^4)} \\
&+ O(\lambda^5)
\end{aligned} \tag{2.95}$$

The purpose of the error analysis is to see the detailed forms of the error terms of order λ^{k+1}. From this analysis it turns out that the error terms are more than 2-body, making it hard to improve the construction by introducing terms of order λ^{k+1} in $\hat{\mathbf{V}}$ to cancel the error terms.

2.7.3 Gap scaling

Now we return to the formulation presented in the beginning of the section where the penalty Hamiltonian $\mathbf{H}$ has a gap Δ, which is a free parameter now (instead of $k-1$ assumed previously). Then according to Equation 2.97, for simulating a target Hamiltonian $\mathbf{H}_{\text{targ}} = \alpha \mathbf{S}_1\mathbf{S}_2 \cdots \mathbf{S}_k$ we let

$$\left(\frac{k-1}{\Delta}\right)^{k-1} \frac{-k(-\lambda)^k}{(k-1)!} = |\alpha| \cdot \text{sgn}(\alpha).$$

Table 2.1
Analytical expressions for $f(\lambda)$ in the example cases. Here we only list k up to 5.

	Target Hamiltonian	$f(\lambda)$
2-body	YY	$-2\lambda^2$
3-body	XYZ	$-\frac{3}{2}\lambda^2$
4-body	$XYZZ$	$-\frac{4}{3}\lambda^2 - \frac{2}{27}\lambda^4$
5-body	$XYZZY$	$-\frac{5}{4}\lambda^2 - \frac{5}{192}\lambda^4$

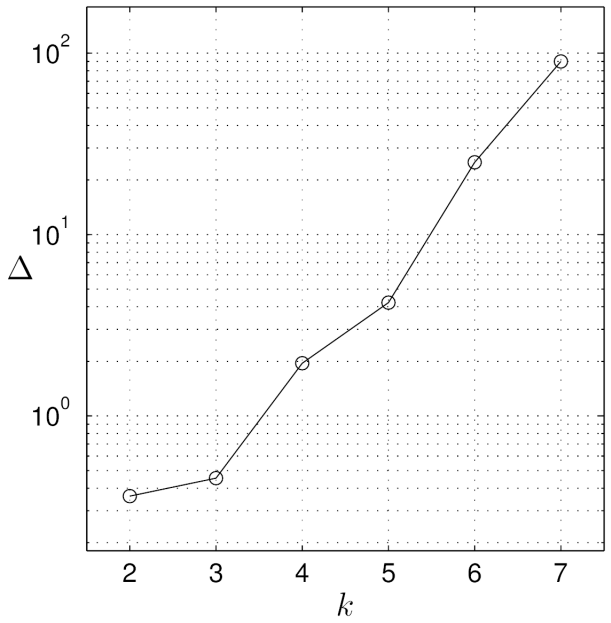


Fig. 2.9. Scaling of the spectral gap Δ as a function of k. Here $\alpha = 0.01$ and $\epsilon = 0.001$. For each case we let $\mathbf{H}_{\text{targ}} = \alpha \mathbf{X}_1 \mathbf{X}_2 \cdots \mathbf{X}_k$. The value of Δ is numerically found as the value that yields the spectral difference between $\widetilde{\mathbf{H}}_{\text{eff}}$ and $\mathbf{H}_{\text{id}}$ being ϵ.

By rearranging the terms we get the expression for λ:

$$\lambda = (-1)^{k+1} \left[\frac{(k-1)!}{k(k-1)^{k-1}} |\alpha| \Delta^{k-1} \right]^{1/k} \text{sgn}(\alpha). \tag{2.96}$$

For a prescribed value of α, there is a value of Δ such that the spectral difference between $\widetilde{\mathbf{H}}_{\text{eff}}$ and $\mathbf{H}_{\text{id}}$ is exactly ϵ. We numerically find such Δ. Figure 2.9 shows the numerically optimized Δ as a function of k. The plots resembles a straight line on a log-linear plot, showing that Δ scales exponentially as a function of k, which provides evidence for the previously established statement that $\Delta = \Omega(\epsilon^{-k})$.

2.7.4 Connection between Bloch formalism and self-energy

Following [9], we have so far used the formulation of perturbation theory due to Bloch [132] to analyze the gadget construction, while the earlier work [7] as well as

several other works on perturbative reductions [1,2,8] employs the standard Feymann-Dyson series formalism. Here we re-consider the construction in [9] in the light of the usual Feynmann-Dyson series and derive connection with the original Bloch formalism. We illustrate the connection with a concrete example. We start by showing how Bloch formalism and the standard perturbation series gives apparently different terms at a fixed order. Then we clarify such difference by a closer examination of the Bloch series.

Consider a target Hamiltonian $\mathbf{H}_{\text{targ}} = \alpha\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4$ and a gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ according to [9]. Here $\mathbf{H}$ is a diagonal Hamiltonian acting on 4 ancilla qubits $w_1, \cdots, w_4$ such that its energy $E_j = j(4-j)$, $j = 0, 1, 2, 3, 4$, for an eigenstate with j ancilla qubits flipped to $|1\rangle$. Let $\mathbf{P}_j$ be a projector onto the eigenspace of energy E_j. Then clearly in this example the ground state projector $\mathbf{P}_0 = \mathbf{P}_4$. Also $\mathbf{P}_1 = \mathbf{P}_3$. The perturbation $\mathbf{V} = \lambda \sum_{i=1}^{4} \boldsymbol{\sigma}_i \otimes \mathbf{X}_{w_i}$.

Analysis in Bloch formalism

According to Bloch formalism, the effective Hamiltonian $\mathbf{H}_{\text{eff}}$ of the lowest $2^4 \times 2$ eigenstates gives

$$\mathbf{H}_{\text{eff}} = \mathcal{U}\mathcal{A}\mathcal{U}^{-1} \tag{2.97}$$

where $\mathcal{U}$ and $\mathcal{A}$ are linear operators each having an expansion that could be derived from Schrödinger equation [9,132]. In particular, $\mathcal{U} = \mathbf{P}_0 + O(\lambda)$ and $\mathcal{U}^{-1} = \mathbf{P}_0 + O(\lambda)$. Let $\mathcal{A} = \mathcal{A}^{(1)} + \mathcal{A}^{(2)} + \mathcal{A}^{(3)} + \mathcal{A}^{(4)} + \cdots$ where each $\mathcal{A}^{(t)}$ is proportional to λ^t. Also define operator $\mathbf{S}^l$ by

$$\mathbf{S}^l = \begin{cases} \displaystyle\sum_{j \neq 0} \frac{\mathbf{P}_j}{(-E_j)^l} & l > 0 \\ -\mathbf{P}_0 & l = 0 \end{cases} \tag{2.98}$$

Then

$$\mathcal{A}^{(1)} + \mathcal{A}^{(2)} = \lambda \mathbf{P}_0 \hat{\mathbf{V}} \mathbf{P}_0 + \lambda^2 \mathbf{P}_0 \hat{\mathbf{V}} \mathbf{S}^1 \hat{\mathbf{V}} \mathbf{P}_0 = -\frac{4}{3}\lambda^2 \mathbf{P}_0. \tag{2.99}$$

Also we can show $\mathcal{A}^{(3)} = \lambda^3(\mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{P}_0 + \mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^2\hat{\mathbf{V}}\mathbf{S}^0\hat{\mathbf{V}}\mathbf{P}_0) = 0$ and

$$\begin{aligned}\mathcal{A}^{(4)} &= \lambda^4(\mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{P}_0 + \mathbf{P}_0\hat{\mathbf{V}}\mathbf{S}^2\hat{\mathbf{V}}\mathbf{S}^0\hat{\mathbf{V}}\mathbf{S}^1\hat{\mathbf{V}}\mathbf{P}_0) \\ &= -\frac{2}{27}\lambda^4\mathbf{P}_0 + \left(-\frac{2}{3}\lambda^4\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4\right)\otimes\mathbf{P}_X\end{aligned} \tag{2.100}$$

where $\mathbf{P}_X = |0000\rangle\langle 1111| + |1111\rangle\langle 0000|$ is a projector acting on the ancilla qubits. The effective Hamiltonian according to Eq. (2.97) becomes

$$\mathbf{H}_{\text{eff}} = f(\lambda)\mathbf{\Pi} + \mathbf{H}_{\text{targ}}\otimes\mathbf{\Pi}_X + O(\lambda^5). \tag{2.101}$$

Here $f(\lambda) = -\frac{4}{3}\lambda^2 - \frac{2}{27}\lambda^4$, $\mathbf{\Pi}$ is the projector onto the lowest $2^4\times 2$ lowest eigenstates of $\widetilde{\mathbf{H}}$, and $\mathbf{\Pi}_X = \mathcal{U}\mathbf{P}_X\mathcal{U}^{-1}$. We choose λ such that $-\frac{2}{3}\lambda^4 = \alpha$ (the case where $\alpha > 0$ could be addressed by a simple modification of $\mathbf{V}$).

Using the Bloch formalism, we can see that when simulating the target Hamiltonian (say, measuring its energy spectrum), we prepare the gadget Hamiltonian, measure the low energy levels of the gadget Hamiltonian and shift the *measured* value by a function $f(\lambda)$ to correct for the actual spectrum of the target Hamiltonian. As we will see in the upcoming discussion, adopting the standard formulation allows us to modify the gadget construction such that the spectral shift $f(\lambda)$ is "incorporated" into the gadget construction and we could directly obtain the energies of the target Hamiltonian by measuring the gadget Hamiltonian.

Analysis using the standard Feynmann-Dyson series

In the standard formalism, the ground state subspace

$$\mathcal{L}_- = \text{span}\{|0000\rangle_{w_1w_2w_3w_4}, |1111\rangle_{w_1w_2w_3w_4}\}$$

and $\mathcal{L}_+ = \text{span}\{|x\rangle_{w_1w_2w_3w_4}, x\in\{0,1\}^4 | h(x) > 0\}$ where $h(\cdot)$ is the Hamming weight of a bit string. Let $\mathbf{\Pi}_\pm$ be projectors onto $\mathcal{L}_\pm$ respectively. Define the operator-valued resolvent (otherwise known as Green's function) as $\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$. Similarly

define $\tilde{\mathbf{G}}(z) = (z\mathbf{I} - \widetilde{\mathbf{H}})^{-1}$. The self-energy expansion $\mathbf{\Sigma}_-(z) = z\mathbf{I} - (\mathbf{\Pi}_-\tilde{\mathbf{G}}(z)\mathbf{\Pi}_-)^{-1}$ then reads

$$\mathbf{\Sigma}_-(z) = \mathbf{V}_- + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-} + \cdots .$$

The k^{th} order term is $\mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{k-2}\mathbf{G}_+\mathbf{V}_{+-}$ for $k \geq 2$. Evaluating the series term by term, we have on the first order $\mathbf{V}_- = 0$, on the second order

$$\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} = -\frac{4}{3}\lambda^2\mathbf{\Pi}_-, \tag{2.102}$$

and the third order being 0 since there is no 3-step transition that goes from all-0 to either all-1 or back to all-0 ancilla states. The fourth order reads

$$\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-} = -\frac{2}{3}\lambda^4\mathbf{\Pi}_- + \left(-\frac{2}{3}\lambda^4\right)\boldsymbol{\sigma}_1\boldsymbol{\sigma}_2\boldsymbol{\sigma}_3\boldsymbol{\sigma}_4 \otimes \mathbf{P}_X. \tag{2.103}$$

Note in Eq. (2.103) that the coefficient of $\mathbf{\Pi}_-$ is $-2/3$ instead of $-2/27$ in Eq. (2.100), with $\mathbf{\Pi}_- \equiv \mathbf{P}_0$ in the previous section on Bloch formalism. We will address this discrepancy by a closer look at the Bloch expansion (2.97).

Bloch expansion: a closer look

In Bloch's original work[3], he defines the low-energy sector of $\widetilde{\mathbf{H}}$ as being spanned by orthonormal wave functions $|\alpha\rangle$. In other words, the low-energy subspace projector $\mathbf{\Pi} = \sum_\alpha |\alpha\rangle\langle\alpha|$. Let $|\alpha_0\rangle = \mathbf{P}_0|\alpha\rangle$. Then $|\alpha_0\rangle$ is in the ground state subspace of $\mathbf{H}$ and is in general *not* normalized. Also let $|\bar{\alpha}_0\rangle$ be a state in the ground state subspace of $\mathbf{H}$ such that $\mathbf{\Pi}|\bar{\alpha}_0\rangle = |\alpha\rangle$. $|\bar{\alpha}_0\rangle$ is also not necessarily normalized. In the light of the above definitions we could see that (see also appendix in [9])

$$\mathbf{P}_0 = \sum_\alpha |\alpha_0\rangle\langle\bar{\alpha}_0|, \quad \mathcal{U} = \sum_\alpha |\alpha\rangle\langle\bar{\alpha}_0|, \quad \mathcal{A} = \sum_\alpha E_\alpha|\alpha_0\rangle\langle\bar{\alpha}_0|. \tag{2.104}$$

To recover the -2/3 coefficient from Eq. (2.97), one needs to expand not only $\mathcal{A}$, but *all* operators $\mathcal{U}$, $\mathcal{U}^{-1}$ and $\mathcal{A}$ and glean terms of order λ^4 from the product.

[3]I thank Stephen Jordan for translating the original paper [132] from French into English, and Ryan Babbush for sharing with me the translated version so that it became accessible to me.

First, with $\mathbf{\Pi} = \mathcal{U}\mathbf{P}_0\mathcal{U}^{-1}$ we obtain that $\mathcal{U}^{-1} = \sum_\alpha |\alpha_0\rangle\langle\alpha|$. We could further show that in fact $\mathcal{U}^{-1} = \mathbf{P}_0\mathbf{\Pi}$. Indeed,

$$\begin{aligned} \mathbf{P}_0\mathbf{\Pi} &= \sum_{\alpha,\beta} |\alpha_0\rangle\langle\bar{\alpha}_0|\beta\rangle\langle\beta| \\ &= \sum_{\alpha,\beta} |\alpha_0\rangle\delta_{\alpha\beta}\langle\beta| \\ &= \sum_{\alpha} |\alpha_0\rangle\langle\alpha| \\ &= \mathcal{U}^{-1} \end{aligned} \tag{2.105}$$

where going from first line to the second we have used $\langle\bar{\alpha}_0|\beta\rangle = (\langle\alpha| + \langle\alpha^\perp|)|\beta\rangle = \langle\alpha|\beta\rangle = \delta_{\alpha\beta}$. Using the expansion of $\mathbf{\Pi}$ from [132], it turns out that

$$\mathcal{U}^{-1} = \mathbf{P}_0 + \mathbf{P}_0\mathbf{V}\mathbf{S}^1 + \mathbf{P}_0\mathbf{V}\mathbf{S}^1\mathbf{V}\mathbf{S}^1 + \mathbf{P}_0\mathbf{V}\mathbf{S}^2\mathbf{V}\mathbf{P}_0 + \mathbf{P}_0\mathbf{V}\mathbf{P}_0\mathbf{V}\mathbf{S}^2 + O(\lambda^3). \tag{2.106}$$

Whatever ends up in the self-energy $\mathbf{\Sigma}_-(z)$ must be in the ground state subspace of $\mathbf{H}$, or in other words

$$\mathbf{\Sigma}_-(z) = \mathbf{P}_0\mathcal{U}\mathcal{A}\mathcal{U}^{-1}\mathbf{P}_0. \tag{2.107}$$

Hence up to $O(\lambda^3)$ the expansion of $\mathcal{U}^{-1}$ is only contributing $\mathbf{P}_0$ and

$$\mathbf{P}_0\mathbf{V}\mathbf{S}^2\mathbf{V}\mathbf{P}_0 = \frac{4}{9}\lambda^2\mathbf{P}_0$$

to the expansion of $\mathbf{\Sigma}_-(z)$. Substituting $\mathcal{U}^{-1} = \mathbf{P}_0 + \frac{4}{9}\lambda^2\mathbf{P}_0$, $\mathcal{U} = \mathbf{P}_0 + O(\lambda)$ and the expansion of $\mathcal{A}$ into Eq. (2.107), gleaning λ^4 one could recover the the coefficient -2/3 from Eq. (2.103).

In conclusion, Eq. (2.107) essentially draws the connection between the two formalisms of perturbation theory. The contributions at each order of λ computed with either formalism should be the same.

2.8 YY gadget

Summary. The gadgets which we have presented so far are intended to reduce the locality of the target Hamiltonian. Here we present another type of gadget, called "creation" gadgets [33], which simulate the type of effective couplings that are not present in the gadget Hamiltonian. Many creation gadgets proposed so far are modifications of existing reduction gadgets. For example, the ZZXX gadget in [33], which is intended to simulate $\mathbf{Z}_i\mathbf{X}_j$ terms using Hamiltonians of the form

$$\mathbf{H}_{ZZXX} = \sum_i \Delta_i \mathbf{X}_i + \sum_i h_i \mathbf{Z}_i + \sum_{i,j} J_{ij} \mathbf{Z}_i \mathbf{Z}_j + \sum_{i,j} K_{ij} \mathbf{X}_i \mathbf{X}_j, \tag{2.108}$$

is essentially a 3- to 2-body gadget with the target term $\mathbf{A} \otimes \mathbf{B} \otimes \mathbf{C}$ being such that the operators $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are $\mathbf{X}$, $\mathbf{Z}$ and identity respectively. Therefore the analyses on 3- to 2- body reduction gadgets that we have presented for finding the lower bound for the gap Δ are also applicable to this ZZXX creation gadget.

Note that YY terms can be easily realized via bases rotation if single-qubit Y terms are present in the Hamiltonian in Equation 2.108. Otherwise it is not *a priori* clear how to realize YY terms using $\mathbf{H}_{ZZXX}$ in Equation 2.108. We will now present the first YY gadget which starts with a universal Hamiltonian of the form Equation 2.108 and simulates the target Hamiltonian $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{Y}_i \mathbf{Y}_j$. The basic idea is to use the identity $\mathbf{X}_i \mathbf{Z}_i = \iota \mathbf{Y}_i$ where $\iota = \sqrt{-1}$ and induce a term of the form $\mathbf{X}_i \mathbf{Z}_i \mathbf{Z}_j \mathbf{X}_j = \mathbf{Y}_i \mathbf{Y}_j$ at the 4^{th} order. Introduce ancilla qubit w and apply a penalty $\mathbf{H} = \Delta |1\rangle\langle 1|_w$. With a perturbation $\mathbf{V}$ we could perform the same perturbative expansion as previously. Given that the 4^{th} order perturbation is $\mathbf{V}_{-+}\mathbf{V}_{+}\mathbf{V}_{+}\mathbf{V}_{+-}$ up to a scaling constant. we could let single $\mathbf{X}_i$ and $\mathbf{X}_j$ be coupled with $\mathbf{X}_w$, which causes both $\mathbf{X}_i$ and $\mathbf{X}_j$ to appear in $\mathbf{V}_{-+}$ and $\mathbf{V}_{+-}$. Furthermore, we couple single $\mathbf{Z}_i$ and $\mathbf{Z}_j$ terms with $\mathbf{Z}_w$. Then $\frac{1}{2}(\mathbf{I}+\mathbf{Z}_w)$ projects single $\mathbf{Z}_i$ and $\mathbf{Z}_j$ onto the + subspace and causes them to appear in $\mathbf{V}_+$. For $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{Y}_1 \mathbf{Y}_2$, the full expressions for the gadget Hamiltonian is the following: the penalty Hamiltonian $\mathbf{H} = \Delta |1\rangle\langle 1|_w$

acts on the ancilla qubit. The perturbation $\mathbf{V} = \mathbf{V}_0 + \mathbf{V}_1 + \mathbf{V}_2$ where $\mathbf{V}_0$, $\mathbf{V}_1$, and $\mathbf{V}_2$ are defined as

$$\begin{aligned} \mathbf{V}_0 &= \mathbf{H}_{\text{else}} + \mu(\mathbf{Z}_1 + \mathbf{Z}_2) \otimes |1\rangle\langle 1|_w + \mu(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2) \otimes \mathbf{X}_w \\ \mathbf{V}_1 &= \frac{2\mu^2}{\Delta}(\mathbf{I} \otimes |0\rangle\langle 0|_w + \mathbf{X}_1\mathbf{X}_2) \\ \mathbf{V}_2 &= -\frac{2\mu^4}{\Delta^3}\mathbf{Z}_1\mathbf{Z}_2 \end{aligned} \tag{2.109}$$

with $\mu = (|\alpha|\Delta^3/4)^{1/4}$. For a specified error tolerance ϵ, we have constructed a YY gadget Hamiltonian of gap scaling $\Delta = O(\epsilon^{-4})$ and the low-lying spectrum of the gadget Hamiltonian captures the spectrum of $\mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ up to error ϵ.

The YY gadget implies that a wider class of Hamiltonians such as

$$\mathbf{H}_{ZZYY} = \sum_i h_i\mathbf{X}_i + \sum_i \Delta_i\mathbf{Z}_i + \sum_{i,j} J_{ij}\mathbf{Z}_i\mathbf{Z}_j + \sum_{i,j} K_{ij}\mathbf{Y}_i\mathbf{Y}_j \tag{2.110}$$

and

$$\mathbf{H}_{XXYY} = \sum_i h_i\mathbf{X}_i + \sum_i \Delta_i\mathbf{Z}_i + \sum_{i,j} J_{ij}\mathbf{X}_i\mathbf{X}_j + \sum_{i,j} K_{ij}\mathbf{Y}_i\mathbf{Y}_j \tag{2.111}$$

can be simulated using the Hamiltonian of the form in Equation 2.108. Therefore using the Hamiltonian in Equation 2.108 one can in principle simulate any finite-norm real valued Hamiltonian on qubits. Although by the QMA-completeness of $\mathbf{H}_{ZZXX}$ one could already simulate such Hamiltonian via suitable embedding, our YY gadget provides a more direct alternative for the simulation.

Analysis. The results in [33] shows that Hamiltonians of the form in Equation 2.108 supports universal adiabatic quantum computation and finding the ground state of such a Hamiltonian is QMA-complete. This form of Hamiltonian is also interesting because of its relevance to experimental implementation [135]. Here we show that with a Hamiltonian of the form in Equation 2.108 we could simulate a target Hamiltonian

$\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \alpha \mathbf{Y}_1 \mathbf{Y}_2$. Introduce an ancilla w and define the penalty Hamiltonian as $\mathbf{H} = \Delta |1\rangle\langle 1|_w$. Let the perturbation $\mathbf{V} = \mathbf{V}_0 + \mathbf{V}_1 + \mathbf{V}_2$ be

$$\begin{aligned} \mathbf{V}_0 &= \mathbf{H}_{\text{else}} + \kappa(\mathbf{Z}_1 + \mathbf{Z}_2) \otimes |1\rangle\langle 1|_w + \kappa(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2) \otimes \mathbf{X}_w \\ \mathbf{V}_1 &= 2\kappa^2 \Delta^{-1}[|0\rangle\langle 0|_w - \text{sgn}(\alpha)\mathbf{X}_1\mathbf{X}_2] \\ \mathbf{V}_2 &= -4\kappa^4 \Delta^{-3} \mathbf{Z}_1 \mathbf{Z}_2. \end{aligned} \tag{2.112}$$

Then the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ is of the form in Equation 2.108. Here we choose the parameter $\kappa = (|\alpha|\Delta^3/4)^{1/4}$. In order to show that the low lying spectrum of $\widetilde{\mathbf{H}}$ captures that of the target Hamiltonian, define $\mathcal{L}_- = \text{span}\{|\psi\rangle \text{ such that } \widetilde{\mathbf{H}}|\psi\rangle = \lambda|\psi\rangle, \lambda < \Delta/2\}$ as the low energy subspace of $\widetilde{\mathbf{H}}$ and $\mathcal{L}_+ \oplus \mathcal{L}_- = \mathcal{H}$. Define $\mathbf{\Pi}_-$ and $\mathbf{\Pi}_+$ as the projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively.

With these notations in place, here we show that the spectrum of $\widetilde{\mathbf{H}}_{<E_*} = \tilde{\mathbf{\Pi}}_- \widetilde{\mathbf{H}} \tilde{\mathbf{\Pi}}_-$ approximates the spectrum of $\mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ with error ϵ. To begin with, the projections of $\mathbf{V}$ into the subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$ can be written as

$$\begin{aligned} \mathbf{V}_- &= \left(\mathbf{H}_{\text{else}} + \underbrace{\frac{\kappa^2}{\Delta}(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2)^2}_{(a)} \underbrace{- \frac{4\kappa^4}{\Delta^3}\mathbf{Z}_1\mathbf{Z}_2}_{(b)} \right) \otimes |0\rangle\langle 0|_w \\ \mathbf{V}_+ &= \left(\mathbf{H}_{\text{else}} + \kappa(\mathbf{Z}_1 + \mathbf{Z}_2) - \frac{2\kappa^2}{\Delta}\text{sgn}(\alpha)\mathbf{X}_1\mathbf{X}_2 - \frac{4\kappa^4}{\Delta^3}\mathbf{Z}_1\mathbf{Z}_2 \right) \otimes |1\rangle\langle 1|_w \\ \mathbf{V}_{-+} &= \kappa(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2) \otimes |0\rangle\langle 1|_w \\ \mathbf{V}_{+-} &= \kappa(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2) \otimes |1\rangle\langle 0|_w. \end{aligned} \tag{2.113}$$

Given the penalty Hamiltonian $\mathbf{H}$, we have the operator valued resolvent $\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$ that satisfies $\mathbf{G}_+(z) = \mathbf{\Pi}_+ \mathbf{G}(z) \mathbf{\Pi}_+ = (z - \Delta)^{-1} |1\rangle\langle 1|_w$. Then the low

lying sector of the gadget Hamiltonian $\widetilde{\mathbf{H}}$ can be approximated by the perturbative expansion Equation 1.26. For our purposes we will consider terms up to the 4^{th} order:

$$\begin{aligned}\mathbf{\Sigma}_-(z) &= \mathbf{V}_- + \frac{1}{z-\Delta}\mathbf{V}_{-+}\mathbf{V}_{+-} + \frac{1}{(z-\Delta)^2}\mathbf{V}_{-+}\mathbf{V}_+\mathbf{V}_{+-} \\ &+ \frac{1}{(z-\Delta)^3}\mathbf{V}_{-+}\mathbf{V}_+\mathbf{V}_+\mathbf{V}_{+-} + \sum_{k=3}^{\infty}\frac{\mathbf{V}_{-+}\mathbf{V}_+^k\mathbf{V}_{+-}}{(z-\Delta)^{k+1}}.\end{aligned} \tag{2.114}$$

Now we explain the perturbative terms that arise at each order. The 1^{st} order is the same as $\mathbf{V}_-$ in Equation 2.113. The 2^{nd} order term gives

$$\frac{1}{z-\Delta}\mathbf{V}_{-+}\mathbf{V}_{+-} = \underbrace{\frac{1}{z-\Delta}\cdot\kappa^2(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2)^2}_{(c)} \otimes |0\rangle\langle 0|_w. \tag{2.115}$$

At the 3^{rd} order, we have

$$\begin{aligned}\frac{1}{(z-\Delta)^2}\mathbf{V}_{-+}\mathbf{V}_+\mathbf{V}_{+-} &= \Bigg(\frac{1}{(z-\Delta)^2}\cdot\kappa^2(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2)\mathbf{H}_{\text{else}}(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2) \\ &+ \underbrace{\frac{1}{(z-\Delta)^2}\frac{4\kappa^4}{\Delta}(\mathbf{X}_1\mathbf{X}_2 - \text{sgn}(\alpha)\mathbf{I})}_{(d)}\Bigg) \otimes |0\rangle\langle 0|_w + O(\Delta^{-1/4}).\end{aligned} \tag{2.116}$$

The 4^{th} order contains the desired YY term:

$$\begin{aligned}\frac{\mathbf{V}_{-+}\mathbf{V}_+\mathbf{V}_+\mathbf{V}_{+-}}{(z-\Delta)^3} &= \Bigg(\underbrace{\frac{1}{(z-\Delta)^3}\cdot 2\kappa^4(\mathbf{X}_1 - \text{sgn}(\alpha)\mathbf{X}_2)^2}_{(e)} - \underbrace{\frac{1}{(z-\Delta)^3}4\kappa^4\mathbf{Z}_1\mathbf{Z}_2}_{(f)} \\ &+ \frac{4\kappa^4\text{sgn}(\alpha)}{(z-\Delta)^3}\mathbf{Y}_1\mathbf{Y}_2\Bigg) \otimes |0\rangle\langle 0|_w + O(\|\mathbf{H}_{\text{else}}\|\cdot\Delta^{-3/4}) \\ &+ O(\|\mathbf{H}_{\text{else}}\|^2\cdot\Delta^{-1/2}).\end{aligned} \tag{2.117}$$

Note that with the choice of $\kappa = (|\alpha|\Delta^3/4)^{1/4}$, all terms of 5^{th} order and higher are of norm $O(\Delta^{-1/4})$. In the 1^{st} order through 4^{th} order perturbations the unwanted terms are labelled as (a) through (f) in Eqs. 2.113, 2.115, 2.116, and 2.117. Note how they compensate in pairs: the sum of (a) and (c) is $O(\Delta^{-1/4})$. The same holds for (d) and (e), (b) and (f). Then the self energy is then

$$\mathbf{\Sigma}_-(z) = (\mathbf{H}_{\text{else}} + \alpha\mathbf{Y}_1\mathbf{Y}_2)\otimes|0\rangle\langle 0|_w + O(\Delta^{-1/4}). \tag{2.118}$$

Let $\Delta = \Theta(\epsilon^{-4})$, then by the Gadget Theorem (1.3.1), the low-lying sector of the gadget Hamiltonian $\widetilde{\mathbf{H}}_{<E_*}$ captures the spectrum of $\mathbf{H}_{\text{targ}} \otimes |0\rangle\langle 0|_w$ up to error ϵ.

The fact that the gadget relies on 4^{th} order perturbation renders the gap scaling relatively larger than it is in the case of subdivision or 3- to 2-body reduction gadgets. However, this does not diminish its usefulness in various applications.

2.9 Conclusion

We have presented improved constructions for the most commonly used gadgets, which in turn implies a reduction in the resources for the many works which employ these current constructions. We presented the first comparison between the known gadget constructions and the first numerical optimizations of gadget parameters. Our analytical results are found to agree with the optimised solutions. The introduction of our gadget which simulates YY-interactions opens many prospects for universal adiabatic quantum computation, particularly the simulation of physics feasible on currently realizable Hamiltonians.

3. PERTURBATIVE GADGETS WITHOUT STRONG INTERACTIONS

In the last chapter we have introduced perturbative gadgets, which are used to construct a quantum Hamiltonian whose low-energy subspace approximates a given quantum k-local Hamiltonian up to an absolute error ϵ. Typically, gadget constructions involve terms with large interaction strengths of order $\text{poly}(\epsilon^{-1})$. In this chapter we present a 2-body gadget construction and prove that it approximates a Hamiltonian of interaction strength $\gamma = O(1)$ up to absolute error $\epsilon \ll \gamma$ using interactions of strength $O(\epsilon)$ instead of the usual inverse polynomial in ϵ. A key component in our proof is a new condition for the convergence of the perturbation series, allowing our gadget construction to be applied in parallel on multiple many-body terms.

We also discuss how to apply this gadget construction for approximating 3- and k-local Hamiltonians. The price we pay for using much weaker interactions is a large overhead in the number of ancillary qubits, and the number of interaction terms per particle, both of which scale as $O(\text{poly}(\epsilon^{-1}))$. Our strong-from-weak gadgets have their primary application in complexity theory (QMA hardness of restricted Hamiltonians, a generalized area law counterexample, gap amplification), but could also motivate practical implementations with several weak interactions simulating a much stronger quantum many-body interaction.

3.1 Overview

The physical properties of (quantum mechanical) spin systems can often be understood in terms of effective interactions arising from the complex interplay of microscopic interactions. Powerful methods for analyzing effective interactions have been developed, for example the renormalization group approach distills effective interac-

tions at different length scales. Another common approach is perturbation theory – treating some interaction terms in the Hamiltonian as a perturbation to a simple original system, giving us a sense of how the fully interacting system behaves. Here, instead of trying to understand an unknown system, we ask an engineering question: how can we build a particular (many-body) effective interaction from local terms of restricted form?

This is where the idea of *perturbative gadgets* provides a powerful answer. Recall from Section 1.6.2 that perturbative gadgets are initially introduced by Kempe, Kitaev and Regev [7] for showing the QMA-hardness of 2-Local Hamiltonian problem and subsequently used and developed further in numerous works [1,8,9,20,33,129,130]. They are convenient tools by which arbitrary many-body effective interactions (which we call the *target Hamiltonian*) can be obtained using a *gadget Hamiltonian* consisting of only two-body interactions. In a broader context, these gadgets have also been used to understand the computational complexity of physical systems (e.g. how hard it is to determine the ground state energy) with restricted geometry of interactions [8], locality [7, 8, 129], or interaction types [33]. Here, we choose to focus on the issue of restricted coupling strengths.

In a nutshell, perturbative gadgets allow us to map between different forms of microscopic Hamiltonians. This is an analogue of how gadgets are used in classical complexity theory, for example in reductions among NP-complete constraint satisfaction problems (e.g. 3-SAT and graph 3-coloring). In the context of combinatorial reductions in classical computation complexity theory, a *gadget* is a finite structure which maps a set of constraints from one optimization problem into a constraint of another problem. Using such gadgets, an instance of 3-SAT (an NP-complete problem) can be efficiently mapped to an instance of graph 3-coloring (also NP-complete [141]). On the other hand, more complex constructions allow us to create more frustrated instances of such problems without significant overhead, resulting in inapproximability as well as the existence of probabilistically checkable proofs [142].

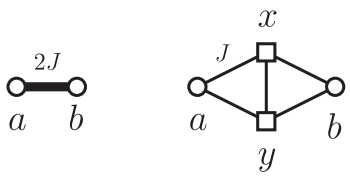


Fig. 3.1. A ferromagnetic interaction $E(a, b) = -2Jab$ of two classical spins $a, b \in \{-1, 1\}$ can be "built" from half-strength interactions involving two extra ancillas. The ground states of the system on the right have $a = b$, while the lowest excited states have $a \neq b$ and energy $4J$ above the ground state energy. Each edge between two classical spins u and v in this illustration represents a term uv in the expression for energy. The ○ nodes symbolize target spins and □ nodes are ancillas.

For classical CSP instances, gadgets can be used to reduce the arity of clauses, to reduce the size of the alphabet, or to reduce the degree of each variable on the constraint graph. Analogously, quantum gadgets [7–9, 126] have been devised for reducing the locality of interactions (analogous to arity reduction in classical CSPs), the dimension of particles (alphabet reduction) and the degree of interaction. These reductions for quantum Hamiltonians give us tools that could help us explore the way to the quantum PCP conjecture [143]. More modestly, gadget translations between types of local Hamiltonians would have implications for the area law [144–146] and other global properties. However, generating approximate *quantum* interactions from a restricted set of terms is not straightforward.

For classical spin systems, creating effective interactions with arbitrary strength by coupling a system to several ancilla degrees of freedom is a relatively simple task. For example, we can create an effective (and twice stronger) ferromagnetic interaction between target spins a, b using two ancilla spins x, y and connecting them to a, b as illustrated in Figure 3.1. The lowest energy states of this new system correspond to the lowest energy states of a system with a ferromagnetic interaction between a and b, with doubled strength.

For general quantum interactions where the target Hamiltonian consists of many-body Pauli operators, the common *perturbative gadget* introduces a strongly bound ancillary system and couples the target spins to it via weaker interactions, treating the latter as a perturbation. The target many-body Hamiltonian is then generated in some low order of perturbation theory of the combined system of ancillary and target spins. Such gadgets first appeared in the proof of QMA-completeness of the 2-local Hamiltonian problem via a reduction from 3-local Hamiltonian [7]. There they helped build effective 3-local interactions from 2-body interactions. Perturbative gadgets can also be used for reducing a target Hamiltonian with general geometry of interactions to a planar interaction graph [8], approximating certain restricted forms of 2-body interactions using other forms of 2-body interactions [32, 33], realizing Hamiltonians exhibiting non-abelian anyonic excitations [147] and reducing k-local interactions to 2-local [9, 129].

For perturbation theory to apply[1], all existing constructions of perturbative gadgets [1, 7–9, 33, 129] require interaction terms or local fields with norm much higher than the strength of the effective interaction which they generate (see Figure 3.2b). However, physically realizable systems often allow only limited spin-spin coupling strengths. The main result of our paper is a way around this problem.

We first build a system with a large spectral gap between the ground state and the first excited state using many relatively weak interactions: consider a collection of n spins that interact with each other (i.e. $O(n^2)$ interaction terms) via ZZ interactions of constant strength J. Then the first excited state of this n-spin system has energy $O(n)$ higher than the ground energy, since the ground state subspace is spanned by $\{|0\rangle^{\otimes n}, |1\rangle^{\otimes n}\}$ and flipping a spin raises the energy by $O(n)$. This way we can use weaker interactions to construct a *core* with a large spectral gap. We then use it to replace the large local field applied onto the single ancilla (Figure 3.2b) with weak interactions of a collection of ancillas (Figure 3.2c). Finally, we connect the

[1] Note that there exist special cases (e.g. Hamiltonians with all terms diagonal in the same basis) when one can analyze the Hamiltonian with non-perturbative techniques [138, 148].

target spins to multiple ancillas instead of just one, which allows us to use weaker β to achieve the same effective interaction strength between the target spins (Figure 3.2d).

Let us review a few definitions and then state our results precisely. An n-qubit Hamiltonian is an $2^n \times 2^n$ Hermitian matrix; it is *k-local* or *k-body* (for a constant k) if it can be written as a sum of $M \leq \text{poly}(n)$ terms $\mathbf{H}_j$, each acting non-trivially on a distinct set of at most k qubits. Furthermore, we require[2] $\|\mathbf{H}_j\| \leq \text{poly}(n)$, and that the entries of $\mathbf{H}_j$ be specified by $\text{poly}(n)$ bits. The smallest eigenvalue of $\mathbf{H}$ is its *ground state energy*, and we denote it $\lambda(\mathbf{H})$. We use $\lambda_j(\mathbf{H})$ to represent the j-th smallest eigenvalue of $\mathbf{H}$, hence $\lambda(\mathbf{H}) = \lambda_1(\mathbf{H})$. Taking a 2-local Hamiltonian acting on n qubits, we can associate it with an *interaction graph* $G(V, E)$. Every vertex $v \in V$ corresponds to a qubit, and there is an edge $e \in E$ between vertices a and b if and only if there is a non-zero 2-local term $\mathbf{H}_e$ on qubits a and b such that $\mathbf{H}_e$ is neither 1-local nor proportional to the identity operator. More generally, we can pair a k-local Hamiltonian with its *interaction hypergraph* in which the k-local terms correspond to hyper-edges involving (at most) k vertices. Note that we depict all 2-local terms on the same spins as a single edge. Next, because we can decompose any 2-local Hamiltonian term in the Pauli basis[3], we can define a *Pauli edge* of an interaction graph G as an edge between vertices a and b associated with an operator $\gamma_{ab}\, \mathbf{P}_a \otimes \mathbf{Q}_b$ where $\mathbf{P}, \mathbf{Q} \in \{\mathbf{I}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ are Pauli matrices and γ_{ab} is a real number signifying the coupling coefficient. We refer to the maximum value of $|\gamma_{ab}|$ as the *interaction strength* of the Hamiltonian. For an interaction graph in which every edge is a *Pauli edge*, the degree of a vertex is called its *Pauli degree*. The maximum Pauli degree of a vertex in an interaction graph is the Pauli degree of the graph.

[2]We use the operator norm $\|\cdot\|$, defined as $\|\mathbf{M}\| \equiv \max_{|\psi\rangle \in \mathcal{M}} |\langle\psi|\mathbf{M}|\psi\rangle|$ for an operator $\mathbf{M}$ acting on a Hilbert space $\mathcal{M}$).

[3]For example, the spin chain Hamiltonian $\mathbf{H} = \frac{1}{2}\sum_{i=1}^{n} |01-10\rangle\langle 01-10|_{i,i+1}$ has interaction edges between successive spins. Each 2-local interaction can be rewritten in the Pauli basis as $\frac{1}{4}(\mathbf{I}\otimes\mathbf{I} - \mathbf{X}\otimes\mathbf{X} - \mathbf{Y}\otimes\mathbf{Y} - \mathbf{Z}\otimes\mathbf{Z})$. It gives us an overall energy shift (from the first term), and three Pauli edges with weight $-\frac{1}{4}$.

We start in Section 3.2 with a theoretical framework of perturbation theory that is used throughout our discussion, and then in Section 3.3 we present and prove our main result – a gadget construction that simulates a target 2-local Hamiltonian using arbitrarily weak 2-local couplings and ancilla particles, summarized in the following Theorem:

Theorem 3.1.1 (Effective 2-body interactions from weak couplings) *Consider the Hamiltonian* $\mathbf{H}_{targ} = \mathbf{H}_{else} + \sum_{j=1}^{M} \gamma_j \, \mathbf{A}_{a_j} \otimes \mathbf{B}_{b_j}$ *on* n *qubits, with* $a_j, b_j \in [n]$ *labeling the qubits that the operators* $\mathbf{A}, \mathbf{B}$ *in the* j^{th} *term act on.* $\mathbf{H}_{targ}$ *consists of*

1. *a Hamiltonian* $\mathbf{H}_{else}$ *with a non-negative spectrum, obeying* $\|\mathbf{H}_{else}\| \leq poly(n)$*, which corresponds to terms in the Hamiltonian that we will* not *decompose into gadgets, and*

2. M *distinct 2-local interaction terms, acting on an* n *qubit system, with an interaction graph of Pauli degree* p*, assuming* $M \leq poly(n)$*, and bounded interaction strength* $\gamma_{max} = \max_j |\gamma_j| = O(1)$*.*

Then for any $\epsilon > 0$ *and* $\epsilon \ll \gamma$*, there exists a Hamiltonian* $\widetilde{\mathbf{H}}$ *which is a sum of* $\mathbf{H}_{else}$ *and a 2-local (gadgetized) Hamiltonian with interaction strengths* $O(\epsilon)$*, whose low-lying spectrum approximates the full spectrum of* $\mathbf{H}_{targ}$ *as* $|\lambda_j(\widetilde{\mathbf{H}}) - \lambda_j(\mathbf{H}_{targ})| \leq \epsilon$ *for all* j *from 1 to* 2^n*. The new Hamiltonian* $\widetilde{\mathbf{H}}$ *acts on* $n + poly(\|\mathbf{H}_{else}\|, \epsilon^{-1}, M)$ *qubits and has an interaction graph of Pauli degree* $poly(p, \|\mathbf{H}_{else}\|, \epsilon^{-1}, M)$*.*

Note that if we want to "gadgetize" the entire target Hamiltonian, $\mathbf{H}_{\text{else}}$ is simply zero. In the remainder of the paper, for $\widetilde{\mathbf{H}}$ and $\mathbf{H}_{\text{targ}}$ in Theorem 3.1.1, when we refer to $\widetilde{\mathbf{H}}$ approximating $\mathbf{H}_{\text{targ}}$ *up to error* ϵ, we mean the following. The low-lying eigenstates of $\widetilde{\mathbf{H}}$ are ϵ-close to $|\phi_j\rangle \otimes |0 \cdots 0\rangle_{anc}$ where $|0 \cdots 0\rangle_{\text{anc}}$ is the state of the ancilla qubits of $\widetilde{\mathbf{H}}$ (the norm of the difference between the vectors is no greater than ϵ), and the low-lying spectrum of $\widetilde{\mathbf{H}}$ is ϵ-close to $\{\lambda_j\}$.

At first glance Theorem 1 seems naively true: we could always consider a given target term $\gamma \mathbf{A} \otimes \mathbf{B}$ as a sum of m identical but smaller terms $\frac{\gamma}{m} \mathbf{A} \otimes \mathbf{B}$ and treat each small term with a separate gadget. Presumably, these gadgets are of weaker interaction strengths than a single gadget applied onto the target term directly. However, if we intend to simulate $\gamma \mathbf{A} \otimes \mathbf{B}$ up to error ϵ, we need to simulate each of the small terms up to error ϵ/m, which would translate into interaction strength in the gadget Hamiltonian scaling as $\text{poly}(\epsilon^{-1})$ regardless. Hence this idea does not improve the interaction strength asymptotically. Our contribution here is to show that we could improve the interaction strength from $\text{poly}(\epsilon^{-1})$ to $\text{poly}(\epsilon)$.

Our main Theorem 3.1.1 deals with 2-local target Hamiltonians, built from 2-local gadgets. What about gadget constructions for reducing 3-local interactions [7, 8, 129] or k-local interactions [9, 129] to 2-local ones? Here we generalize Theorem 1 to propose gadget constructions for 3- and k-body target Hamiltonians. In particular:

Corollary 1 (3-body terms from weak 2-body interactions) *Let us consider a Hamiltonian* $\mathbf{H}_{targ} = \mathbf{H}_{else} + \sum_{i=1}^{M} \gamma_i \, \mathbf{A}_{a_i} \otimes \mathbf{B}_{b_i} \otimes \mathbf{F}_{f_i}$, *with Here* $a_j, b_j, f_j \in [n]$ *labeling the qubits that the operators* $\mathbf{A}, \mathbf{B}, \mathbf{F}$ *in the* j^{th} *term act on. Here* $\mathbf{H}_{targ}$ *consists of*

1. $\mathbf{H}_{else}$ *(the part we will not decompose into gadgets), a Hamiltonian with a non-negative spectrum, satisfying* $\|H_{else}\| \leq poly(n)$, *and*

2. *a sum of* M *interaction terms that are 3-local, acting on an* n *qubit system, with an interaction graph of Pauli degree* p *and ground state energy* $\lambda(\mathbf{H}_{targ})$, *assuming* $M \leq poly(n)$. *The interaction strength of* $\mathbf{H}_{targ}$ *satisfies* $\gamma_{max} = \max_j |\gamma_j| = O(1)$.

Then for any choice of $\epsilon > 0$, *there exists a Hamiltonian* $\widetilde{\mathbf{H}}$ *that consists of* $\mathbf{H}_{else}$ *(the part we leave intact) and a sum of* M *terms that are 2-local, with interaction strength* $O(\epsilon)$, *acting on a system with* $n + poly(\|\mathbf{H}_{else}\|, \epsilon^{-1}, M)$ *qubits, with an interaction graph of Pauli degree* $poly(p, \|\mathbf{H}_{else}\|, \epsilon^{-1}, M)$ *and* $|\lambda_j(\widetilde{\mathbf{H}}) - \lambda_j(\mathbf{H}_{targ})| \leq \epsilon$ *for all* j.

We outline the proof of Corollary 1 in Section 3.4. An important property of these new constructions is that they can be repeated in parallel, in essence generating arbitrary strong interactions from weak ones. Thus, we can effectively rescale interaction strengths and amplify the eigenvalue gap of a local Hamiltonian. The price we pay is the addition of many ancillas and a large increase in the number of interactions per particle.

Whereas Theorem 3.1.1 states that a 2-local target Hamiltonian can be gadgetized to a Hamiltonian with arbitrarily weak interactions, Corollary 1 states that the same could be accomplished for a 3-local target Hamiltonian. (In Section 3.4 we also generalize it to k-local Hamiltonians.)

Next, besides producing a gadget Hamiltonian with weak interactions that generates the target Hamiltonian, we could also generate the target Hamiltonian multiplied by a positive factor θ. In case where $\theta > 1$, this can be viewed as a coupling strength amplification relative to the original target k-local Hamiltonian (see Corollary 2 below). The basic proof idea is to view the rescaled target Hamiltonian $\theta\mathbf{H}$ (with $\theta > 1$) as a sum of $O(\theta)$ copies of itself with interaction strength $O(1)$. Using the gadget constructions from [9], we transform the k-local Hamiltonian $\theta\mathbf{H}$ to a 2-local one. Finally, using our 2-body gadget construction in this work, we translate this Hamiltonian to one with only weak interactions (2-body).

Corollary 2 (Coupling strength amplification by gadgets) *Let* $\mathbf{H} = \sum_{j=1}^{M} \mathbf{H}_j$ *be a* k*-local Hamiltonian on* n *qubits where* $M = poly(n)$ *and each* $\mathbf{H}_j$ *satisfies* $\|\mathbf{H}_j\| \leq s$ *for some constant* s*. Let* $|\phi_j\rangle$ *and* λ_j *be the* j*-th eigenstate and eigenvalue of* $\mathbf{H}$*. Choose a magnifying factor* $\theta > 1$ *and an error tolerance* $\epsilon > 0$*. Then there exists a 2-local Hamiltonian* $\widetilde{\mathbf{H}}$ *with interactions of strength* $O(1)$ *or weaker. The low-lying*

eigenstates of $\widetilde{\mathbf{H}}$ *are* ϵ*-close*[4] *to* $|\phi_j\rangle \otimes |0\cdots 0\rangle_{anc}$ *where* $|0\cdots 0\rangle_{anc}$ *is the state of the ancilla qubits of* $\widetilde{\mathbf{H}}$*, and the low-lying spectrum of* $\widetilde{\mathbf{H}}$ *is* ϵ*-close to* $\theta\{\lambda_j\}$.

What is the efficiency of this way of amplifying the couplings? If we do it in a series of reductions from k to $\lceil k/2 \rceil$ to $\lceil \lceil k/2 \rceil /2 \rceil$, etc., to 2-body interactions, the final gadget Hamiltonian will act on a system whose total number of qubits scales exponentially in k (which of course is not a problem for $k = 3$).

3.2 Effective interactions based on perturbation theory

The purpose of a perturbative gadget is to approximate a target n-qubit Hamiltonian $\mathbf{H}_{\text{targ}}$ by a gadget Hamiltonian $\widetilde{\mathbf{H}}$ which uses a restricted form of interactions among the n qubits that $\mathbf{H}_{\text{targ}}$ acts on and poly(n) additional ancilla qubits. The subspace spanned by the lowest 2^n eigenstates of $\widetilde{\mathbf{H}}$ should approximate the spectrum of $\mathbf{H}_{\text{targ}}$ up to a prescribed error tolerance ϵ in the sense that the j-th lowest eigenvalue of $\widetilde{\mathbf{H}}$ differs from that of $\mathbf{H}_{\text{targ}}$ by at most ϵ and the inner product between the corresponding eigenstates of $\widetilde{\mathbf{H}}$ and $\mathbf{H}_{\text{targ}}$ (assume no degeneracy) is at least $1 - \epsilon$. These error bounds can be established using perturbation theory [7, 8]. There are various versions of perturbation theory available for constructing and analyzing gadgets (for a review see [149]). For example, Jordan and Farhi [9] use Bloch's formalism, while Bravyi et al. rely on the Schrieffer-Wolff transformation [129]. For the gadgets in Sec. 3.3, we use the technique from [7, 8].

Let us now review the basic ideas underlying the construction of effective Hamiltonians from gadgets. The gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ is a sum of an unperturbed Hamiltonian $\mathbf{H}$ and a perturbation $\mathbf{V}$. $\mathbf{H}$ acts only on the ancilla space, energetically penalizing certain configurations, and favoring a specific ancilla state or subspace. Second, we have a perturbation $\mathbf{V}$ describing how the target spins interact with the ancillas.

[4] By ϵ*-close* we mean the norm of the difference between the two quantities (scalar, vector or matrix operator) is $\leq \epsilon$.

Let us introduce the following notations: let λ_j and $|\psi_j\rangle$ be the j^{th} eigenvalue and eigenvector of $\mathbf{H}$ and similarly define $\tilde{\lambda}_j$ and $|\tilde{\psi}_j\rangle$ for $\widetilde{\mathbf{H}}$, assuming all the eigenvalues are labeled in a weakly increasing order ($\lambda_1 \leq \lambda_2 \leq \cdots$, similarly for $\tilde{\lambda}_j$). Using a cutoff value λ_*, let us call $\mathcal{L}_- = \text{span}\{|\psi_j\rangle : \lambda_j \leq \lambda_*\}$ the *low-energy subspace* and $\mathcal{L}_+ = \text{span}\{|\psi_j\rangle : \lambda_j > \lambda_*\}$ the *high-energy subspace*. Let $\mathbf{\Pi}_-$ and $\mathbf{\Pi}_+$ be the orthogonal projectors onto the subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$. For an operator $\mathbf{O}$ we define the partitioning of $\mathbf{O}$ into these subspaces as $\mathbf{O}_- = \mathbf{\Pi}_-\mathbf{O}\mathbf{\Pi}_-$, $\mathbf{O}_+ = \mathbf{\Pi}_+\mathbf{O}\mathbf{\Pi}_+$, $\mathbf{O}_{-+} = \mathbf{\Pi}_-\mathbf{O}\mathbf{\Pi}_+$ and $\mathbf{O}_{+-} = \mathbf{\Pi}_+\mathbf{O}\mathbf{\Pi}_-$. We define similar notations $\tilde{\mathcal{L}}_-$ and $\tilde{\mathcal{L}}_+$ for $\widetilde{\mathbf{H}}$.

Our first goal is to understand $\widetilde{\mathbf{H}}|_{\tilde{\mathcal{L}}_-}$, the restriction of the gadget Hamiltonian to its low-energy subspace. Let us consider the operator-valued *resolvent* $\tilde{\mathbf{G}}(z) = (z\mathbf{I} - \tilde{\mathbf{H}})^{-1}$ where $\mathbf{I}$ is the identity operator. Similarly let us define $\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$. Note that $\tilde{\mathbf{G}}^{-1}(z) - \mathbf{G}^{-1}(z) = -\mathbf{V}$, which allows an expansion of $\tilde{\mathbf{G}}$ in powers of $\mathbf{V}$:

$$\tilde{\mathbf{G}} = (\mathbf{G}^{-1} - \mathbf{V})^{-1} = \mathbf{G}(\mathbf{I} - \mathbf{V}\mathbf{G})^{-1} = \mathbf{G} + \mathbf{G}\mathbf{V}\mathbf{G} + \mathbf{G}\mathbf{V}\mathbf{G}\mathbf{V}\mathbf{G} + \cdots . \tag{3.1}$$

It is also standard to define the *self-energy* $\mathbf{\Sigma}_-(z) = z\mathbf{I} - (\tilde{\mathbf{G}}_-(z))^{-1}$. It is important because the spectrum of $\mathbf{\Sigma}_-(z)$ gives an approximation to the spectrum of $\widetilde{\mathbf{H}}_-$, since by definition $\widetilde{\mathbf{H}}_- = z\mathbf{I} - \mathbf{\Pi}_-(\tilde{\mathbf{G}}^{-1}(z))\mathbf{\Pi}_-$ while $\mathbf{\Sigma}_-(z) = z\mathbf{I} - (\mathbf{\Pi}_-\tilde{\mathbf{G}}(z)\mathbf{\Pi}_-)^{-1}$. As explained in [8], if $\mathbf{\Sigma}_-(z)$ is roughly constant in some range of z (see Theorem 3.2.1 below for details) then $\mathbf{\Sigma}_-(z)$ is (loosely speaking) playing the role of $\widetilde{\mathbf{H}}_-$. This was formalized in Theorem 3 in [7] (and improved in Theorem A.1 in [8]). Similarly to [8], we choose to work with $\mathbf{H}$ whose lowest eigenvalue is zero and whose spectral gap is Δ. In [7], the gadget theorem (Theorem 3) is proven by establishing a sequence of Lemmas. Out of these, Lemma 5 requires the condition $\|\mathbf{V}\| < \frac{\Delta}{2}$, with the consequence being the separation of subspaces, namely $\tilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}$. Therefore, we here remove the condition $\|\mathbf{V}\| < \frac{\Delta}{2}$ and use $\tilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}$ as an alternative assumption, giving us a slightly modified Gadget approximation theorem:

Theorem 3.2.1 (Gadget approximation theorem, modified from [7]) *Let* $\mathbf{H}$ *be a Hamiltonian with a gap* Δ *between its ground state and first excited state. As-*

suming the ground state energy of $\mathbf{H}$ *is 0, let* $\lambda_* = \Delta/2$. *Consider a bounded norm perturbation* $\mathbf{V}$. *The perturbed Hamiltonian is then* $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$. *Following the notations introduced previously, if the following holds:*

1. $\tilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}$, *with* $\mathcal{L}_+ = span\{|\psi_j\rangle : \lambda_j \leq \lambda_*\}$ *for* $|\psi_j\rangle$ *eigenvectors of* $\mathbf{H}$ *and* $\tilde{\mathcal{L}}_- = span\{|\tilde{\psi}_j\rangle : \tilde{\lambda}_j \leq \lambda_*\}$ *for* $|\tilde{\psi}_j\rangle$ *eigenvectors of* $\widetilde{\mathbf{H}}$.

2. *There is an effective Hamiltonian* $\mathbf{H}_{\mathit{eff}}$ *with a spectrum contained in* $[E_1, E_2]$ *for some* $\epsilon > 0$ *and* $E_1 < E_2 < \Delta/2 - \epsilon$, *such that for every* $z \in [E_1 - \epsilon, E_2 + \epsilon]$, *the self-energy* $\mathbf{\Sigma}_-(z)$ *obeys* $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\mathit{eff}}\| \leq \epsilon$.

then all the eigenvalues of $\widetilde{\mathbf{H}}_-$ *are close to the eigenvalues of* $\mathbf{H}$, *obeying*

$$|\lambda_j(\mathbf{H}_{\mathit{eff}}) - \lambda_j(\widetilde{\mathbf{H}}_-)| \leq \epsilon.$$

The first, subspace condition, says that a combination of the unperturbed high-energy eigenstates of $\mathbf{H}$ can not by themselves form a low-energy state of $\widetilde{\mathbf{H}}$. We choose to avoid the original stronger condition $\|\mathbf{V}\| \leq \frac{\Delta}{2}$ from [7], since it imposes limitations on the *global* properties of the gadget construction, in particular the number of ancillas we use, disregarding the structure of the perturbation. One might question whether the use of perturbation theory is sensible if we assume that the perturbation $\|\mathbf{V}\|$ is no longer necessarily small compared to the spectral gap Δ (we want to use a large number M of gadgets). However, such use has been justified previously by Bravyi et al. [129] in a similar context.

To apply Theorem 3.2.1, a series expansion for the self-energy $\mathbf{\Sigma}_-(z) = z\mathbf{I} - \tilde{\mathbf{G}}_-^{-1}(z)$ is truncated at some low order, for which $\mathbf{H}_{\text{eff}}$ is approximated. Using the series expansion of $\tilde{G}$ in (3.1), the self-energy can be expanded as (see [7] for details)

$$\mathbf{\Sigma}_-(z) = \mathbf{H}_- + \mathbf{V}_- + \mathbf{V}_{-+}\mathbf{G}_+(z)\mathbf{V}_{+-} + \mathbf{V}_{-+}\mathbf{G}_+(z)\mathbf{V}_+\mathbf{G}_+(z)\mathbf{V}_{+-} + \cdots, \quad (3.2)$$

with $\mathbf{G}_+(z) = \mathbf{\Pi}_+(z\mathbf{I} - \mathbf{H})^{-1}\mathbf{\Pi}_+$. The 2^{nd} and higher order terms in this expansion give rise to effective many-body interactions. Introducing auxiliary spins and a suitable selection of 2-local $\mathbf{H}$ and $\mathbf{V}$, we can engineer $\mathbf{\Sigma}_-(z)$ to be ϵ-close to

$\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes \mathbf{\Pi}_-$ (here $\mathbf{\Pi}_-$ is the projector to the ground state subspace of the ancillas) in the range of z considered in Theorem 3.2.1. Therefore with $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\| \leq \epsilon$, condition 2 of Theorem 3.2.1 is satisfied.

In the next Section, we will look at the usual 2-body gadgets and see how the second order terms in the self-energy result in the desired effective Hamiltonian. Then we present our construction that involves more ancillas with weaker interactions, and show that the effective Hamiltonian is again what we want, and that the conditions for Theorem 3.2.1 are met.

3.3 A new gadget for 2-body interactions

We can decompose any 2-local interaction of spin-$\frac{1}{2}$ particles in the Pauli basis[5], using terms of the form $\gamma \mathbf{A} \otimes \mathbf{B}$, with the operator $\mathbf{A}$ acting on spin a and $\mathbf{B}$ acting on spin b, and γ the *interaction strength*. Without loss of generality, we can also use $\pm$ Pauli matrices, and fix the coupling strengths to be positive. It will be enough to show how to replace any such "Pauli" interaction in our system by a gadget, aiming at the target interaction $\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \gamma \mathbf{A} \otimes \mathbf{B}$, with $\mathbf{H}_{\text{else}}$ some $O(1)$-norm, 2-local Hamiltonian. First, we briefly review the existing constructions [1, 8, 129] for generating $\mathbf{H}_{\text{targ}}$ using a gadget Hamiltonian $\widetilde{\mathbf{H}}$. Then we present a new 2-body gadget which simulates an arbitrary $\gamma = O(1)$ strength 2-local interaction using a gadget Hamiltonian with terms of strength only $o(1)$, "building" quantum interactions from many weaker ones.

The usual construction. Consider a target 2-local term involving two qubits a, b as depicted in Fig. 3.2(a). The standard construction of a gadget Hamiltonian $\widetilde{\mathbf{H}}$ that captures the 2-local target term is shown in Fig. 3.2(b). First, we introduce an ancilla qubit w bound by a local field, with the Hamiltonian $\mathbf{H} = -\frac{\Delta}{2}\mathbf{Z}_w$. Alternatively, up to a spectral shift we could write $\mathbf{H} = \Delta|1\rangle\langle 1|_w$ where $|1\rangle\langle 1|_w = \frac{1}{2}(\mathbf{I} - \mathbf{Z}_w)$. Then

[5] It is useful that the Pauli matrices $\mathbf{A}, \mathbf{B} \in \{\mathbf{I}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ square to identity, because $\mathbf{A}^2$ and $\mathbf{B}^2$ terms in our effective Hamiltonian will become simple overall energy shifts

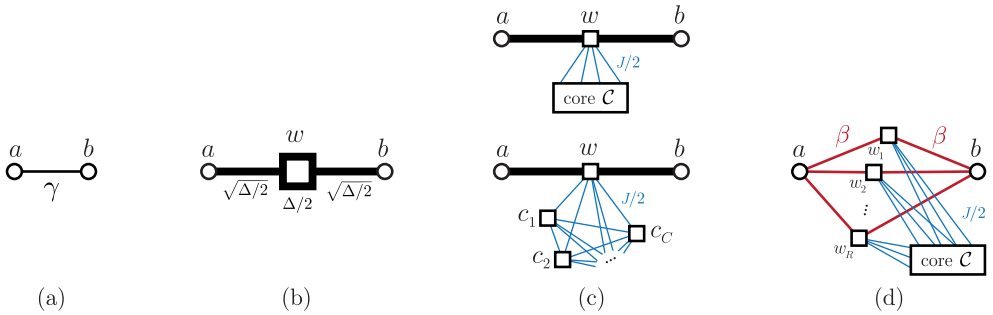


Fig. 3.2. Interaction graphs for effective two-body interaction mediated by ancilla qubits. Each node represents a particle. The size of the node indicates the strength of local field applied onto it. The width of each edge shows the strength of the interaction between the particles that the edge connects. (a): The desired 2-local interaction between target spins a, b. (b): The usual perturbative gadget uses a single ancilla w in a strong local field, and large-norm interactions with the target spins. (c): We can replace the strong local field $\Delta/2$ by ferromagnetic interactions with a fixed *core* – a group of C "core" ancilla qubits located in a field of strength $J/2$, interacting with each other ferromagnetically (as a complete graph), with strength $J/2$. (d): Instead of the strong interactions between target spins a, b and a single ancilla w, we can use R different "direct" ancillas (labeled as w_1, w_2, $\cdots$, w_R) and weaker interactions of strength β.

we let w interact with a and b through $\sqrt{\Delta/2}\,\mathbf{A}\otimes\mathbf{I}\otimes\mathbf{X}_w$ and $-\sqrt{\Delta/2}\,\mathbf{I}\otimes\mathbf{B}\otimes\mathbf{X}_w$, and choose $\Delta=\Theta(\epsilon^{-1})$. We can view these terms as a perturbation to $\mathbf{H}$. The low energy effective Hamiltonian calculated from (3.2) is approximately $\mathbf{A}\otimes\mathbf{B}\otimes|0\rangle\langle 0|_w$ (up to an overall energy shift) [1]. Here "up to an error ϵ" means that the j-th lowest eigenvalue of $\widetilde{\mathbf{H}}$ differs from that of $\mathbf{H}_{\text{targ}}$ by at most ϵ and the inner product between the corresponding eigenstates of $\widetilde{\mathbf{H}}$ and $\mathbf{H}_{\text{targ}}$ (assume no degeneracy) is at least $1-\epsilon$.

Our construction. In the usual construction, with better precision (decreasing ϵ), the spectral gap Δ (related to local field strength) and interaction strengths grow as *inverse* polynomials in ϵ. We now suggest a 2-body gadget which simulates an arbitrary $O(1)$ strength target interaction using a gadget Hamiltonian of only $O(\epsilon)$ interaction strength *i.e.* without the need for large-norm terms. We build it in a sequence of steps illustrated in Fig. 3.2.

The first step is to reduce the large local field Δ in Fig. 3.2(b). Let us call the ancilla w directly interacting with the target spins a *direct* ancilla. We add a *core* $\mathcal{C}$ – a set of C ancilla qubits $c_1,\ldots,c_C$, with a complete graph of ferromagnetic (ZZ) interactions of strength $\frac{J}{2}$, and in a local field of strength $\frac{J}{2}$ where $J=O(\epsilon)$. We then let the direct ancilla w interact (ferromagnetically) with each of the core ancillas, as in Fig. 3.2(c). The Hamiltonian for the direct and core ancillas then reads

$$\frac{J}{2}\sum_{c\in\mathcal{C}}(\mathbf{I}-\mathbf{Z}_w\mathbf{Z}_c)+\underbrace{\frac{J}{2}\sum_{c\in\mathcal{C}}(\mathbf{I}-\mathbf{Z}_c)+\frac{J}{2}\sum_{c,c'\in\mathcal{C}}(\mathbf{I}-\mathbf{Z}_c\mathbf{Z}_{c'})}_{\equiv\mathbf{H}_{\mathcal{C}}}\,. \tag{3.3}$$

$\mathbf{H}_{\mathcal{C}}$ is the Hamiltonian describing the core $\mathcal{C}$. The ground state of this Hamiltonian is $|0\rangle_w|0\cdots 0\rangle_{\mathcal{C}}$, and the gap between its ground and first excited state $|1\rangle_w|0\cdots 0\rangle_{\mathcal{C}}$ is $\Delta=JC$. Here C is the number of ancillas in the core $\mathcal{C}$.

The second step is to use R direct ancillas $w_1, \ldots, w_R$ instead of just one, connecting each of them to the core ancillas as in Figure 3.2(d). The Hamiltonian then becomes

$$\mathbf{H} = \frac{J}{2} \sum_{i=1}^{R} \sum_{c \in \mathcal{C}} (\mathbf{I} - \mathbf{Z}_{w_i} \mathbf{Z}_c) + \mathbf{H}_{\mathcal{C}}. \tag{3.4}$$

Its ground state is $|0 \cdots 0\rangle_w \otimes |0 \cdots 0\rangle_{\mathcal{C}}$ (here we use the subscript w to refer to all the direct ancillas connected to the target qubits), and the gap between the two lowest energies is still $\Delta = JC$.

We want to engineer an effective interaction $\mathbf{H}_{\text{targ}} = \gamma \mathbf{A}_a \otimes \mathbf{B}_b + \mathbf{H}_{\text{else}}$, where the first term is our desired Hamiltonian, and $\mathbf{H}_{\text{else}}$ is a finite-norm Hamiltonian that includes all the other terms that we want to leave unchanged by this gadget. Starting with the Hamiltonian $\mathbf{H}$ (3.4), we add a perturbation

$$\mathbf{V} = \mathbf{H}_{\text{else}} + \beta \sum_{i=1}^{R} \left(\mathbf{A}_a \otimes \mathbf{X}_{w_i} - \mathbf{B}_b \otimes \mathbf{X}_{w_i} \right), \tag{3.5}$$

where $\beta > 0$ is the strength of the interactions between the target spins and the direct ancillas. Showing that we can use perturbation theory to obtain the effective Hamiltonian crucially relies on Theorem 3.2.1, and we will justify that its conditions hold later. Let us now prepare the notations and tools for this. Let $\mathcal{L}_-$ be the subspace with the ancillas in the state $|0\rangle^{\otimes(R+C)}$. Denote $\mathcal{L}_+$ the subspace orthogonal to $\mathcal{L}_-$ and let Π_- and Π_+ be the projectors onto these subspaces. We then have

$$\mathbf{V}_- = \mathbf{\Pi}_- \mathbf{V} \mathbf{\Pi}_- = \mathbf{H}_{\text{else}} \otimes \mathbf{\Pi}_-, \tag{3.6}$$

$$\mathbf{V}_{-+} = \mathbf{\Pi}_- \mathbf{V} \mathbf{\Pi}_+ = \beta (\mathbf{A}_a - \mathbf{B}_b) \otimes \sum_{i=1}^{R} |0\rangle\langle 1|_{w_i}, \tag{3.7}$$

$$\mathbf{V}_{+-} = \mathbf{\Pi}_+ \mathbf{V} \mathbf{\Pi}_- = \beta (\mathbf{A}_a - \mathbf{B}_b) \otimes \sum_{i=1}^{R} |1\rangle\langle 0|_{w_i}, \tag{3.8}$$

$$\mathbf{V}_+ = \mathbf{\Pi}_+ \mathbf{V} \mathbf{\Pi}_+ = \mathbf{H}_{\text{else}} \otimes \mathbf{\Pi}_+ + \beta \sum_{i=1}^{R} (\mathbf{A}_a - \mathbf{B}_b) \otimes \mathbf{\Pi}_+ \mathbf{X}_{w_i} \mathbf{\Pi}_+. \tag{3.9}$$

The low-energy sector of the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ can be described by the self-energy expansion (3.2). Let us compute the terms up to the second order.

- At the 0th order, $\mathbf{H}_- = 0$ by definition.
- At the 1st order, $\mathbf{V}_-$ is given by (3.6).
- At the 2nd order, we have the term $\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-}$, where $\mathbf{V}_{-+}$ and $\mathbf{V}_{+-}$ can be computed from (3.7) and (3.8). We also need the operator-valued resolvent

$$\mathbf{G}_+(z) = \sum_{x:h(x)>0} \frac{1}{z - h(x)\Delta} |x\rangle\langle x|.$$

  A second order transition process from the low energy subspace back to itself can only take the form $|0\rangle^{\otimes R} \to |x\rangle \to |0\rangle^{\otimes R}$, with x an R-bit string of Hamming weight 1 (there are R qubits that can be flipped there and back). Hence, the only non-trivial terms in the product $\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-}$ have the form $\beta(\mathbf{A}_a - \mathbf{B}_b) \cdot \frac{1}{z-\Delta} \cdot \beta(\mathbf{A}_a - \mathbf{B}_b)$. Altogether, we have R of these terms, so the second order term becomes $\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} = \frac{1}{z-\Delta} R\beta^2(\mathbf{A}_a - \mathbf{B}_b)^2$.

We could obtain the higher order terms in a similar fashion. In the end, the self-energy expansion becomes

$$\mathbf{\Sigma}_-(z) = \underbrace{\mathbf{H}_{\text{else}}}_{1^{\text{st}}\text{ order}} + \underbrace{\frac{1}{z-\Delta} R\beta^2(\mathbf{A}_a - \mathbf{B}_b)^2}_{2^{\text{nd}}\text{ order}} + \underbrace{\sum_{m=1}^{\infty} \mathbf{V}_{-+}\mathbf{G}_+(\mathbf{V}_+\mathbf{G}_+)^m\mathbf{V}_{+-}}_{\text{error term}}. \tag{3.10}$$

Recall that $G(z) = (z\mathbf{I} - \mathbf{H})^{-1}$. The range of z we consider is $|z| \le \|\mathbf{H}_{\text{else}}\| + |\gamma|$. We can assume $\gamma > 0$ in $\mathbf{H}_{\text{targ}}$ (e.g. by absorbing a possible minus sign into the $\mathbf{A}$ matrix), and choose

$$\beta = \sqrt{\frac{\gamma\Delta}{2R}} = \sqrt{\frac{\gamma JC}{2R}}. \tag{3.11}$$

Since $z \ll \Delta$, we can write $\frac{1}{z-\Delta} = -\frac{1}{\Delta}\left(1 - \frac{z}{\Delta}\right)^{-1} \approx -\frac{1}{\Delta} + O\left(\frac{1}{\Delta^2}\right)$. Then the 1^{st} and 2^{nd} order terms are approximately equal to the desired effective Hamiltonian $\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes \mathbf{\Pi}_-$ up to an overall spectral shift (because $\mathbf{A}^2 = \mathbf{B}^2 = \mathbf{I}$). We will show later (Claim 2) that with good choices of R and C we can make β and J as small as we want.

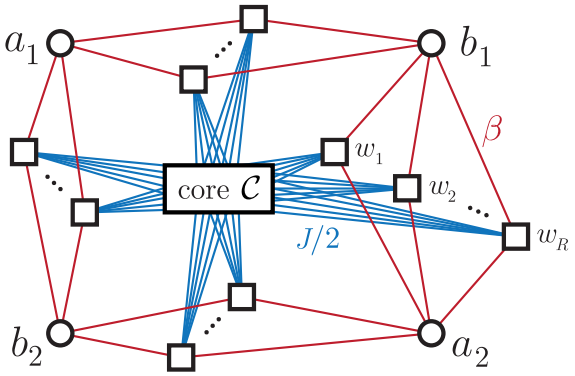


Fig. 3.3. Parallel composition of M (here $M = 4$) two-body gadgets from Fig. 3.2(d), using a single common core with C "core" ancillas. Each gadget has R "direct" ancillas interacting with the target spins. The total number of ancillas is thus $MR + C$.

Parallel 2-body gadgets. So far, we have focused on a single 2-local term in our target Hamiltonian (see Fig. 3.2). Similarly to [8], we can apply our gadgets *in parallel*, which enables us to deal with a target Hamiltonian with M such 2-local terms. Let us then consider a target Hamiltonian of the form

$$\mathbf{H}_{\text{targ}} = \mathbf{H}_{\text{else}} + \sum_{j=1}^{M} \gamma_j \mathbf{A}_{a_j} \otimes \mathbf{B}_{b_j} \tag{3.12}$$

and apply our construction to every term $\gamma_j \mathbf{A}_{a_j} \otimes \mathbf{B}_{b_j}$ in parallel, as in Fig. 3.3. Note that we save a lot of resources by using only a single core. Each target term $\gamma_j \mathbf{A}_{a_j} \otimes \mathbf{B}_{b_j}$ is associated with R direct ancilla qubits $w_1^{(i)}$, $w_2^{(i)}$, $\cdots$, $w_R^{(i)}$ that are connected to target spins a_i and b_i. All of the direct ancillas also interact with each of the C core ancillas. As before, the core consists of C qubits that are fully connected with ferromagnetic (ZZ) interactions of strength $\frac{J}{2}$ and also with local fields of strength $\frac{J}{2}$ on each qubit. Hence the full gadget Hamiltonian for the general 2-local target Hamiltonian in (3.12) takes the form $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ with

$$\begin{aligned} \mathbf{H} &= \frac{J}{2} \sum_{j=1}^{M} \sum_{i=1}^{R} \sum_{c \in \mathcal{C}} (\mathbf{I} - \mathbf{Z}_{w_i^{(j)}} \mathbf{Z}_c) + \mathbf{H}_{\mathcal{C}}, \\ \mathbf{V} &= \mathbf{H}_{\text{else}} + \sum_{j=1}^{M} \beta_j \sum_{i=1}^{R} \left(\mathbf{A}_{a_j} - \mathbf{B}_{b_j} \right) \otimes \mathbf{X}_{w_i^{(j)}}. \end{aligned} \tag{3.13}$$

where $\mathbf{H}_{\mathcal{C}}$ is the core Hamiltonian from (3.4), $\beta_j = \sqrt{\frac{\gamma_j JC}{2R}}$ and the spectral gap between the ground state and the first excited state of $\mathbf{H}$ is $\Delta = JC$. Computing the self-energy expansion as in (3.10) for the gadget Hamiltonian in (3.13) yields a contribution $-\frac{1}{z-\Delta} \sum_{j=1}^{M} \beta_j^2 R (\mathbf{A}_{a_j} - \mathbf{B}_{b_j})^2$ at the second order (see Claim 2 for more details). Because each term in the perturbative expansion $\mathbf{\Sigma}_-(z)$ corresponds to a sequence of state transitions from $\mathcal{L}_-$ to $\mathcal{L}_+$ and back[6], the second order contribution comes from those transitions where one ancilla is flipped from $|0\rangle$ to $|1\rangle$ and back to

[6]Note that in fact $\mathcal{L}_- = \text{span}\{|0\cdots 0\rangle_w |0\cdots 0\rangle_{\mathcal{C}}\}$ where the subscript w refers to all the ancillas $w_1^{(j)}$, $w_2^{(j)}$, $\cdots$, $w_R^{(j)}$, for $j = 1, 2, \cdots, M$. The transitions that contribute to the perturbative expansion $\mathbf{\Sigma}_-(z)$ are restricted to only to the direct ancillas $|0\cdots 0\rangle_w$ since the core ancillas do not interact with the target qubits.

$|0\rangle$. Such transitions cannot involve more than one ancilla qubit. Hence we can regard the second order transitions involving different ancillas as occurring independently of each other (in parallel). This enables the 2-body gadgets to capture multiple 2-local target terms, and is much more effective than a "serial" approach: constructing a gadget for a single 2-body interaction, calling what we get $\mathbf{H}_{\text{else}}$, then building another gadget for another 2-body interaction, and so on.

In order to show that the low-lying subspace of our gadget Hamiltonian $\widetilde{\mathbf{H}}$ captures the spectrum of $\mathbf{H}_{\text{targ}}$ using Theorem 3.2.1, it is necessary to establish that $\widetilde{\mathbf{H}}$ meets both conditions of the theorem. The first condition, $\tilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}$, requires the vectors the unperturbed high-energy states not to become perturbed low energy states by themselves. We will prove this as Claim 1 below. The second condition says that the self-energy expansion $\mathbf{\Sigma}_-(z)$ can be approximated by an effective Hamiltonian when z is in a certain range. We establish this as Claim 2 for $\widetilde{\mathbf{H}}$ by proving that the perturbation series converges for $\mathbf{\Sigma}_-(z)$. Theorem 3.1.1 then follows from Theorem 3.2.1 with $\widetilde{\mathbf{H}}$ being the Hamiltonian in (3.13).

3.3.1 The 2-local construction satisfies the subspace condition.

The first condition in Theorem 3.2.1 is a property of the high-energy subspace of the original Hamiltonian. We need it in order to avoid the need to bound the norm of the whole perturbation. Let us provide a high-level description of the condition and the ideas behind its proof.

Consider the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ defined in (3.13). We need to lower bound the lowest energy $E_+ = \min_\psi \langle\psi|\widetilde{\mathbf{H}}|\psi\rangle$ of a state $|\psi\rangle$ that comes from the subspace $\mathcal{L}_+$, the excited subspace of $\mathbf{H}$, spanned by states orthogonal to the state $|0\cdots0\rangle_w$. The terms in $\mathbf{H}$ involve only ancilla qubits, while $\mathbf{V}$ includes $\mathbf{H}_{\text{else}}$, and terms that couple some computational (target) qubit a and a direct ancilla w. These 2-local terms have form $\beta_w \mathbf{A}_a \otimes \mathbf{X}_w$, with interaction strengths $|\beta_w| \leq \beta_{\max} = O(1)$, as in Figure 3.4(a). We now want to show that E_+ is strictly above $\lambda_* = \frac{\Delta}{2}$. To do

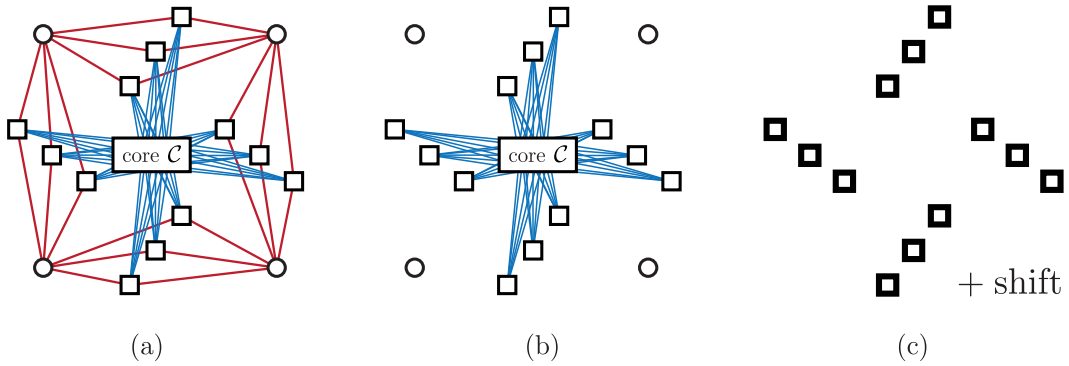


Fig. 3.4. A sequence of gadget Hamiltonians with progressively lower lower bounds on E_+. (a): Taking the terms acting on the target spins to be all the same. (b): Decoupling the target spins from the direct ancillas using $-\mathbf{I}$ operators on the target spins and (weighted) X-fields on the direct ancillas. (c): Replacing the interactions with core ancillas by an overall shift, and a (weighted) Z-field on the direct ancillas, arriving at (3.30).

this, we find a sequence of successively lower lower bound E_+ using a sequence of progressively simpler Hamiltonians, finally arriving at 1-local ones in (3.24), (3.28) and (3.30).

First, we will show that E_+ for the general Hamiltonian $\widetilde{\mathbf{H}}$ is greater or equal to the value of E_+ for a similar system in which all of the operators $\mathbf{A}_a$ are the same (so that they do not compete against each other in lowering the energy) as in Figure 3.4(a). Second, we can only lower E_+ by making all of the operators $\mathbf{A}_a$ identities, and using only operators $-|\beta_w|\,\mathbf{X}_w$ on the direct ancillas. Because the target spins are now independent from the ancillas, the contribution from $\mathbf{H}_{\text{else}}$ is then no larger in magnitude than $\|\mathbf{H}_{\text{else}}\|$. This is depicted in Figure 3.4(b).

We are now left with a Hamiltonian which is a sum of $\mathbf{H}_{\text{else}}$, single-qubit terms on the direct ancilla qubits, and their interactions[7] with the core ancillas. The Hilbert

[7]If the values of β are different for different terms, we still use a single core with a fixed J, fixed C, fixed $\Delta = JC$, and adjust each β_w for each target interaction individually so that the resulting effective interaction strength $\beta_w^2 R/\Delta = O(1)$ is what we desire.

space divides into a direct sum of invariant subspaces labeled by the state of the core ancillas. These subspaces are decoupled (the original Hamiltonian $\mathbf{H}$ and the perturbation $\mathbf{V}$ do not flip the core ancillas), so we can analyze them one by one. We do so for the subspaces with $a \geq 1$ core ancillas flipped to $|1\rangle$, and then finally for the subspace with all core ancillas equal to $|0\rangle$. It turns out that in each such subspace we can map the terms $\mathbf{Z}_w\mathbf{Z}_c, \mathbf{Z}_c\mathbf{Z}_{c'}, \mathbf{Z}_c$ and $\mathbf{H}_{\text{else}}$ of the Hamiltonian[8] to one that is simply an overall shift, and a $-\frac{\Delta_a}{2}\mathbf{Z}_w$ term on each of the direct ancillas, with Δ_a a function of how many ancillas were flipped. The resulting 1-local Hamiltonian illustrated in Figure 3.4(c) can be analyzed, and yields the desired lower bound on E_+. Let us then state and prove our first Claim.

Claim 1 *Consider the 2-body gadget Hamiltonian* $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ *from* (3.13)*, corresponding to a target Hamiltonian* $\mathbf{H}_{targ} = \mathbf{H}_{else} + \sum_{i=1}^{M} \gamma_i \mathbf{A}_i \otimes \mathbf{B}_i$ *with* $\gamma_j \leq O(1)$ *and* $\mathbf{H}_{else}$ *positive semi-definite. Let* Δ *be the spectral gap between the ground and the first excited subspace of* $\mathbf{H}$*, and define a cutoff* $\lambda_* = \Delta/2$*. Following Section 3.2, we define* $\mathcal{L}_+ = span\{|\psi_j\rangle : \lambda_j > \lambda_*\}$ *for* $|\psi_j\rangle$ *eigenvectors of* $\mathbf{H}$*, and* $\tilde{\mathcal{L}}_- = span\{|\tilde{\psi}_j\rangle : \tilde{\lambda}_j < \lambda_*\}$*, for* $|\tilde{\psi}_j\rangle$ *eigenvectors of* $\widetilde{\mathbf{H}}$*. Then if* $\Delta \geq 160M\gamma_{max}$*, with* $\gamma_{max} = \max_{j=1,\cdots,M} |\gamma_j|$*, we have*

$$\tilde{\mathcal{L}}_- \cap \mathcal{L}_+ = \{0\}.$$

We start the proof by exhibiting a sequence of Hamiltonians with progressively lower E_+, and then showing Claim 1 for the last of them.

Let $|\psi\rangle \in \mathcal{L}_+$ be the state with minimum energy for the perturbed Hamiltonian $\widetilde{\mathbf{H}}$, and let us label this minimum energy $E_+ = \langle\psi|\widetilde{\mathbf{H}}|\psi\rangle$. The Hamiltonian $\widetilde{\mathbf{H}}$ connects target spins to direct ancillas via terms of the type $\mathbf{A}_a \otimes \mathbf{X}_j$. We now argue that E_+ can be only lowered if we decouple the target spins from the direct ancillas, and simply use $-\mathbf{I} \otimes \mathbf{X}_j$ instead of $\mathbf{A}_a \otimes \mathbf{X}_j$.

The expectation value $\langle\psi|\widetilde{\mathbf{H}}|\psi\rangle = E_{\mathbf{H}} + E_{\mathbf{V}}$ comes from the expectation value of $\mathbf{H}$ which is diagonal in the computational basis (the Z and ZZ terms involving the

[8]There is no $\mathbf{Z}_w$ term on the direct ancillas, so that a single direct ancilla flip increases the energy by $\Delta = JC$.

ancillas) and the expectation of $\mathbf{V}$, which includes the interactions with target spins as well as the term $\mathbf{H}_{\text{else}}$. Let us rewrite the state $|\psi\rangle$ as

$$|\psi\rangle = \sum_w c_w |w\rangle \otimes |\phi_w\rangle, \tag{3.14}$$

where w is a binary string labeling computational basis state of all the ancillas. The expectation value of the term $\mathbf{H}$ depends only on the magnitudes of the c_w's. The contribution from $\mathbf{H}_{\text{else}}$ is

$$\sum_w |c_w|^2 \langle \phi_w | \mathbf{H}_{\text{else}} | \phi_w \rangle. \tag{3.15}$$

Finally, each term in $\mathbf{V}$ of the form $\mathbf{A}_a \otimes \mathbf{X}_j$ contributes

$$c_v^* c_{v'} \langle v | \mathbf{X}_j | v' \rangle \langle \phi_v | \mathbf{A}_a | \phi_{v'} \rangle \tag{3.16}$$

for every pair v, v' that differ only at bit j. This expression can be positive or negative, depending on c_v and $c_{v'}$. More crucially, its magnitude will depend on $\langle \phi_v | \mathbf{A}_a | \phi_{v'} \rangle$. Because $\mathbf{A}_a$ is a Pauli operator, this magnitude can never exceed 1. Let us now consider a state

$$|\psi'\rangle = \left(\sum_w |c_w| \, |w\rangle \right) \otimes |\phi\rangle, \tag{3.17}$$

with positive coefficients $|c_w|$, and a particular state $|\phi\rangle$ chosen to minimize $\langle \phi | \mathbf{H}_{\text{else}} | \phi \rangle$. The expectation value of $\mathbf{H}$ does not change, while the contribution from $\mathbf{H}_{\text{else}}$ can only decrease, because we have chosen $|\phi\rangle$ to minimize it. In other words, $\langle \psi' | \mathbf{H} | \psi' \rangle \leq \langle \psi | \mathbf{H} | \psi \rangle$ and $\langle \psi' | \mathbf{H}_{\text{else}} | \psi' \rangle \leq \langle \psi | \mathbf{H}_{\text{else}} | \psi \rangle$. Finally, the expectation value of the interaction terms in $\mathbf{V}'$ (when we set $A_a = -\mathbf{I}$) like (3.16) now become

$$-|c_v| \cdot |c_{v'}| \langle v | \mathbf{X}_u | v' \rangle \leq -c_v^* c_{v'} \langle v | \mathbf{X}_u | v' \rangle \langle \phi_v | \mathbf{A}_a | \phi_{v'} \rangle. \tag{3.18}$$

Thus, $\langle \psi' | \mathbf{V}' | \psi' \rangle \leq \langle \psi | \mathbf{V} | \psi \rangle$ and we conclude that the new minimum energy of $\tilde{\mathbf{H}}'$ restricted to $\mathcal{L}_+$ is $E'_+ \leq \langle \psi' | \mathbf{H} | \psi' \rangle + \langle \psi' | \mathbf{V}' | \psi' \rangle \leq E_+$. It means that when we replace the Hamiltonian $\mathbf{V}$ with one that has no interactions between the direct ancillas and the target spins, and uses operators $-\mathbf{X}_w$ on the direct ancillas, E_+ decreases (or remains what it was).

Therefore, we can assume without loss of generality that the Hamiltonian $\widetilde{\mathbf{H}}$ has this special form. We will continue the proof by showing that if $|\psi\rangle \in \mathcal{L}_+$ then $\langle\psi|\widetilde{\mathbf{H}}|\psi\rangle > \lambda_*$ for any normalized state $|\psi\rangle$.

The subspace $\mathcal{L}_+$ is spanned by (direct + core) ancilla qubit states with at least one $|1\rangle$. Let $\mathcal{K}_- = |0\cdots 0\rangle_w$ be the all-zero state the direct ancillas, and let $\mathcal{S}_a = \text{span}\{|x\rangle_\mathcal{C} : h(x) = a\}$ be the subspace of the core ancillas with exactly a qubits[9] in the state $|1\rangle$. Thus, the subspace $\mathcal{L}_+$ splits into two parts as $\mathcal{L}_+ = \mathcal{L}_1 \oplus \mathcal{L}_2$, where[10]

$$\mathcal{L}_1 = \mathcal{H}_w \otimes \left(\bigoplus_{a=1}^{C} \mathcal{S}_a\right), \qquad \mathcal{L}_2 = \mathcal{K}_-^\perp \otimes \mathcal{S}_0. \tag{3.19}$$

The first part $\mathcal{L}_1$ spanned by all the states where the core has at least one qubit $|1\rangle$, while the second part $\mathcal{L}_2$ is spanned by all the states with the core ancillas all $|0\rangle$, and at least one direct ancillas being $|1\rangle$. We now first show that $\forall|\psi\rangle \in \mathcal{L}_1$, $\langle\psi|\tilde{H}|\psi\rangle > \lambda_*$, and then similarly for $\mathcal{L}_2$.

(1) *If* $|\psi\rangle \in \mathcal{L}_1$, *then* $\langle\psi|\widetilde{\mathbf{H}}|\psi\rangle > \lambda_*$.

Let us first consider $|\psi_a\rangle \in \mathcal{H}_w \otimes \mathcal{S}_a$ for some fixed $a \in \{1, 2, \cdots, C\}$. Then $|\psi_a\rangle$ is a (linear combination of) state(s) where a ancillas in the core are $|1\rangle$ and the other $C-a$ core ancillas are $|0\rangle$. We will find a lower bound for $\tilde{E}_{+,a} = \langle\psi_a|\widetilde{\mathbf{H}}|\psi_a\rangle$ by considering each component of $\widetilde{\mathbf{H}}$. Recall from (3.3) the definition of the core Hamiltonian

$$H_\mathcal{C} = \frac{J}{2}\sum_{c\in\mathcal{C}}(\mathbf{I} - \mathbf{Z}_c) + \frac{J}{2}\sum_{c,c'\in\mathcal{C}}(\mathbf{I} - \mathbf{Z}_c\mathbf{Z}_{c'}). \tag{3.20}$$

Then the energy of $|\psi_a\rangle$ with respect to the core Hamiltonian is $E_{\mathcal{C},a} = \langle\psi_a|\mathbf{H}_\mathcal{C}|\psi_a\rangle = Ja(C-a+1) \geq JC$. Let

$$\mathbf{H}_w = \frac{J}{2}\sum_{j=1}^{M}\sum_{i=1}^{R}\sum_{c\in\mathcal{C}}(\mathbf{I} - \mathbf{Z}_{w_i^{(j)}}\mathbf{Z}_c) = \frac{J}{2}\sum_{j=1}^{M}\sum_{i=1}^{R}\left(C\mathbf{I} - \mathbf{Z}_{w_i^{(j)}}\sum_{c\in\mathcal{C}}\mathbf{Z}_c\right) \tag{3.21}$$

[9]Here $h(x)$ is the Hamming weight of the binary string x.
[10]Here $\mathcal{H}_w$ is the Hilbert space of the direct ancillas.

be the interaction Hamiltonian between the direct ancillas and the core ancillas. Recall from (3.13) that $\mathbf{H} = \mathbf{H}_w + \mathbf{H}_\mathcal{C}$. The second equality in (3.21) indicates that $\mathbf{H}_w$ consists of a sum of terms of the form $C\mathbf{I} - \mathbf{Z}_{w_i^{(j)}} \sum_{c\in\mathcal{C}} \mathbf{Z}_c$. Let us focus on such a term for a particular direct ancilla w. Consider the states $|0\rangle_w \otimes |a\rangle_\mathcal{C}$ and $|1\rangle_w \otimes |a\rangle_\mathcal{C}$ with $|a\rangle_\mathcal{C} \in \mathcal{S}_a$ and look at the term $\mathbf{Z}_w \sum_{c\in\mathcal{C}} \mathbf{Z}_c$. Its expectation value in these states is $C - 2a$ and $2a - C$, regardless of the state of the core ancillas. Thus, we get an effective Hamiltonian

$$\mathbf{h}'_w = C\mathbf{I} - (C - 2a)\mathbf{Z}_w \tag{3.22}$$

for each direct ancilla w. Collecting these effective Hamiltonians for each direct ancilla, we get

$$\mathbf{H}'_w = \frac{J}{2} \sum_{j=1}^{M} \sum_{i=1}^{R} \mathbf{h}'_{w_i^{(j)}} = \frac{J}{2} \sum_{k=1}^{N} \mathbf{h}'_k, \tag{3.23}$$

whose lowest energy in the subspace $\mathcal{H}_w \otimes \mathcal{S}_a$ is equal to that of $\mathbf{H}_w$. For convenience, we relabel the direct ancillas by $k = 1, \ldots N$ with $N = MR$ (we are simulating M two-body interactions using R direct ancillas per interaction), and replace the sum over i and j with a single index summation over k.

Let us now add the perturbation $\mathbf{V}$ (3.13). For each direct ancilla k there is a term in $\mathbf{V}$ of the form $\mathbf{v}_k = \beta_k(\mathbf{A} \otimes \mathbf{X}_k - \mathbf{B} \otimes \mathbf{X}_k) = \beta_k \mathbf{O}_{AB} \otimes \mathbf{X}_k$, and we have shown by a sequence of reductions that the lowest energy of $\mathbf{v}_k$ in $\mathcal{H}_k \otimes \mathcal{S}_a$ is lower bounded by that of $\mathbf{v}'_k = 2\beta_k \mathbf{X}_k$. Thus, when we label $\mathbf{V}' = \sum_{k=1}^{N} \mathbf{v}'_k = \sum_{k=1}^{N} 2\beta_k \mathbf{X}_k$, we get a 1-local Hamiltonian

$$\begin{aligned} \widetilde{\mathbf{H}}' &= E_{\mathcal{C},a}\mathbf{I} + \mathbf{H}'_w + \mathbf{V}' = E_{\mathcal{C},a}\mathbf{I} + \sum_{k=1}^{N} \left(2\beta_k \mathbf{X}_k + \frac{JC}{2}\mathbf{I} - \frac{J(C-2a)}{2}\mathbf{Z}_k \right) \\ &= \left(Ja(C - a + 1) + \frac{JCN}{2} \right) \mathbf{I} + \sum_{k=1}^{N} \sqrt{4\beta_k^2 + \frac{J^2}{4}(C-2a)^2}\mathbf{P}_k \end{aligned} \tag{3.24}$$

acting only on the direct ancillas, which gives us a lower bound on E_+, i.e. for any $|\psi_a\rangle \in \mathcal{H}_w \otimes \mathcal{S}_a$,

$$\min_{\mathcal{H}_w \otimes \mathcal{S}_a} \langle \psi_a | \widetilde{\mathbf{H}} | \psi_a \rangle \geq \min_{\mathcal{H}_w \otimes \mathcal{S}_a} \langle \psi_a | \widetilde{\mathbf{H}}' | \psi_a \rangle, \tag{3.25}$$

with $\mathbf{P}_k$ a single qubit operator of the form $\hat{p}\cdot\vec{\hat{\boldsymbol{\sigma}}}$, with $\vec{\hat{\boldsymbol{\sigma}}} = \{\mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ and unit vector $\hat{\mathbf{p}}$. Note that the lower bound (3.25) does not include $\mathbf{H}_{\text{else}}$ in $\widetilde{\mathbf{H}}'$; because $\mathbf{H}_{\text{else}} \geq 0$, we are only lowering the right side by omitting it.

Note that the above argument can be generalized to $\mathcal{L}_1 = \mathcal{H}_w \otimes (\oplus_{a=1}^{C} \mathcal{S}_a)$. For a general $|\psi\rangle \in \mathcal{L}_1$, $|\psi\rangle$ must take the form

$$|\psi\rangle = |\phi\rangle_w \otimes \sum_{a=1}^{C} \eta_a |a\rangle_{\mathcal{C}}, \qquad \text{where} \qquad |a\rangle = \sum_{h(x)=a} c_{a,x}|x\rangle, \quad x \in \{0,1\}^C \tag{3.26}$$

for some sets of complex coefficients $\{\eta_a\}$ and $\{c_{a,x}\}$ that are both normalized. Then $\langle\psi|\mathbf{H}_{\mathcal{C}}|\psi\rangle = \sum_{a=1}^{C} |\eta_a|^2 Ja(C-a+1)$. Let $\mathcal{A}$ be the set of a for which $\eta_a \neq 0$. Let $a_{\max}$ be the value of a in $\mathcal{A}$ that maximizes $(C-2a)^2$. Define

$$\begin{aligned} |\psi'\rangle &= |\phi\rangle_w \otimes |a_{\max}\rangle_{\mathcal{C}}, \\ \mathbf{h}'_{w,a_{\max}} &= C\mathbf{I} - (C-2a_{\max})\mathbf{Z}_w, \\ \mathbf{H}'_{w,a_{\max}} &= \frac{J}{2}\sum_{k=1}^{N} \mathbf{h}'_k. \end{aligned} \tag{3.27}$$

Then $\langle\psi|\mathbf{H}_w|\psi\rangle \geq \langle\psi_a|\mathbf{H}'_{w,a_{\max}}|\psi_a\rangle$ for any $|\phi\rangle_w \in \mathcal{H}_w$. Since the generalization from $\mathcal{H}_w \otimes \mathcal{S}_a$ to $\mathcal{H}_w \otimes (\oplus_{a=1}^{C} \mathcal{S}_a)$ does not concern the direct ancillas, we can use the same argument as before to construct a 1-local Hamiltonian

$$\widetilde{\mathbf{H}}'_{a_{\max}} = \sum_{a=1}^{C} |\eta_a|^2 \left(Ja(C-a+1) + \frac{JCN}{2} \right) \mathbf{I} + \sum_{k=1}^{N} \sqrt{4\beta_k^2 + \frac{J^2}{4}(C-2a_{\max})^2}\, \mathbf{P}_{k,a_{\max}} \tag{3.28}$$

such that for any $|\psi_a\rangle \in \mathcal{L}_1$, there always exists a value $a_{\max}$ such that

$$\min_{|\psi\rangle\in\mathcal{L}_1} \langle\psi|\widetilde{\mathbf{H}}|\psi\rangle \geq \min_{|\psi\rangle\in\mathcal{L}_1} \langle\psi|\widetilde{\mathbf{H}}'_{a_{\max}}|\psi\rangle.$$

Let us now find a lower bound on $\langle\psi|\widetilde{\mathbf{H}}'_{a_{\max}}|\psi\rangle$. Note that $Ja(C-a+1) \geq JC$ for any $a = 1, 2, \cdots, C$. Let $\beta_{\max} = \max_{k=1,2,\cdots,N} |\beta_k|$. Noting that $\mathbf{P}_{k,a_{\max}}$ in (3.28) is a unit-norm operator, for any $|\psi\rangle \in \mathcal{L}_1$ we get

$$\langle\psi|\widetilde{\mathbf{H}}'_{a_{\max}}|\psi\rangle \geq \left(JC + \frac{JCN}{2} \right) - N\sqrt{4\beta_{\max}^2 + \frac{J^2}{4}(C-2a_{\max})^2}$$

$$
\begin{aligned}
&= JC + \frac{JCN}{2} - \frac{JCN}{2}\sqrt{1 + \frac{16\beta_{\max}^2}{J^2C^2}} \\
&\geq JC - \frac{JCN}{2} \cdot \frac{16\beta_{\max}^2}{2J^2C^2} = JC - \frac{4N\beta_{\max}^2}{JC} \\
&= \Delta - \frac{4MR\beta_{\max}^2}{\Delta} \geq \frac{79\Delta}{80} > \frac{\Delta}{2} = \lambda_*, \qquad (3.29)
\end{aligned}
$$

where we have used $2R\beta_{\max}^2/\Delta = \gamma_{\max}$ from (3.11) and asked for $\Delta \geq 160M\gamma_{\max}$ in the last line. Here $\gamma_{\max} = \max_{j=1,\cdots,M} |\gamma_j|$ where γ_j are coefficients in the target Hamiltonian. Putting (3.29) into (3.25), we get $E_+ > \frac{\Delta}{2} = \lambda_*$. We have thus shown the desired lower bound on E_+ in the subspace $\mathcal{L}_1$. Let us now deal with the other part, $\mathcal{L}_2$.

(2) *If* $|\psi\rangle \in \mathcal{L}_2$, *then* $\langle\psi|\widetilde{\mathbf{H}}|\psi\rangle > \lambda_*$.

Any state in the subspace $\mathcal{L}_2 = \mathcal{K}_-^{\perp} \otimes \mathcal{S}_0$ has the core ancillas in the state $|0\cdots0\rangle_{\mathcal{C}}$, hence $\langle\psi|\mathbf{H}_{\mathcal{C}}|\psi\rangle = 0$. To find a lower bound for the energy of $\mathbf{H}_w$ in this subspace, we use the construction $\mathbf{H}'_w$ in (3.23) with $a = 0$. For the energy of $\mathbf{V}$ we use the same simplifying argument and obtain (again) a 1-local Hamiltonian acting only on the $N = MR$ direct ancillas (cf. Equation 3.24)

$$
\widetilde{\mathbf{H}}'_0 = \sum_{k=1}^{N} \left(\frac{\Delta}{2}\mathbf{I} - \frac{\Delta}{2}\mathbf{Z}_k - 2\beta_k \mathbf{X}_k \right) = \sum_{k=1}^{N} \mathbf{S}_k, \qquad (3.30)
$$

such that

$$
\min_{|\psi\rangle \in \mathcal{L}_2} \langle\psi|\widetilde{\mathbf{H}}|\psi\rangle \geq \min_{|\psi\rangle \in \mathcal{L}_2} \langle\psi|\widetilde{\mathbf{H}}'_0|\psi\rangle. \qquad (3.31)
$$

We now show that the energy of any direct ancilla state orthogonal to

$$
\mathcal{K}_- = \text{span}\{|0\ldots0\rangle_w\}
$$

is strictly lower bounded by $\lambda_* = \Delta/2$. Since the core ancilla state will always be $|0\cdots0\rangle_{\mathcal{C}}$, we will exclude it from our discussion and thus omit the w subscript for the direct ancilla. All quantum states in the proof from here on refer to the state of the direct ancillas.

To show the energy lower bound we use induction on the number of direct ancillas, n. Let

$$E_n = \min_{|\phi\rangle \perp |0\rangle^{\otimes n}} \langle\phi| \sum_{k=1}^{n} \mathbf{S}_k |\phi\rangle. \tag{3.32}$$

Specifically, we prove the following statement:

$$E_n \geq \frac{3\Delta}{4} - \delta_n, \qquad \text{with} \qquad \delta_n = \frac{40n\beta^2}{9\Delta}, \qquad n = 1, \cdots, N. \tag{3.33}$$

We start with the initial case $n = 1$. There the only state orthogonal to $|0\rangle$ is $|1\rangle$. Hence $E_1 = \Delta$, which satisfies (3.33). Now assume (3.33) holds for some n. An $(n+1)$-qubit state that is orthogonal to $|0\cdots0\rangle$ (denoted by the superscript $\oslash$) must have the form

$$|\psi_{n+1}^{\oslash}\rangle = a|\xi_n^{\oslash}\rangle|0\rangle + b|\phi_n^{\oslash}\rangle|1\rangle + c|0\cdots0\rangle|1\rangle, \tag{3.34}$$

where $|\xi_n^{\oslash}\rangle$ and $|\phi_n^{\oslash}\rangle$ are some states that are orthogonal to $|0\cdots0\rangle$. Let us calculate the energy of the state (3.34).

$$\begin{aligned}
E_{n+1} &= \sum_{i=1}^{n} \langle\psi_{n+1}^{\oslash}|\mathbf{S}_i|\psi_{n+1}^{\oslash}\rangle + \langle\psi_{n+1}^{\oslash}|\mathbf{S}_{n+1}|\psi_{n+1}^{\oslash}\rangle \\
&= |a|^2 \sum_{i=1}^{n} \langle\xi_n^{\oslash}|\mathbf{S}_i|\xi_n^{\oslash}\rangle\langle 0|0\rangle + |b|^2 \sum_{i=1}^{n} \langle\phi_n^{\oslash}|\mathbf{S}_i|\phi_n^{\oslash}\rangle\langle 1|1\rangle \\
&\qquad + |c|^2 \sum_{i=1}^{n} \langle 0\cdots0|\mathbf{S}_i|0\cdots0\rangle\langle 0|0\rangle \\
&\quad + 2\mathrm{Re}\left(ab^* \sum_{i=1}^{n} \langle\xi_n^{\oslash}|\mathbf{S}_i|\phi_n^{\oslash}\rangle\langle 0|1\rangle + ac^* \sum_{i=1}^{n} \langle\xi_n^{\oslash}|\mathbf{S}_i|0\cdots0\rangle\langle 0|1\rangle \right. \\
&\qquad \left. + bc^* \sum_{i=1}^{n} \langle\phi_n^{\oslash}|\mathbf{S}_i|0\cdots0\rangle\langle 1|1\rangle \right) \\
&\quad + |a|^2\langle 0|\mathbf{S}|0\rangle_{n+1} + |b|^2\langle 1|\mathbf{S}|1\rangle_{n+1} + |c|^2\langle 1|\mathbf{S}|1\rangle_{n+1} \\
&\quad + 2\mathrm{Re}\left(ab^*\langle\xi_n^{\oslash}|\phi_n^{\oslash}\rangle\langle 0|\mathbf{S}|1\rangle + ac^*\langle\xi_n^{\oslash}|0\cdots0\rangle\langle 0|\mathbf{S}|1\rangle + bc^*\langle\phi_n^{\oslash}|0\cdots0\rangle\langle 1|\mathbf{S}|1\rangle \right).
\end{aligned} \tag{3.35}$$

Note that $\langle 0|1\rangle = 0$ and $\langle\psi_n^{\oslash}|0\cdots0\rangle = \langle\phi_n^{\oslash}|0\cdots0\rangle = 0$. Also recall that $\langle 0|\mathbf{S}_i|0\rangle = 0$, $\langle 1|\mathbf{S}_i|1\rangle = \Delta$ and $\langle 0|\mathbf{S}_i|1\rangle = -2\beta_i$. Hence,

$$E_{n+1} = |a|^2 \sum_{i=1}^{n} \langle\xi_n^{\oslash}|\mathbf{S}_i|\xi_n^{\oslash}\rangle + |b|^2 \sum_{i=1}^{n} \langle\phi_n^{\oslash}|\mathbf{S}_i|\phi_n^{\oslash}\rangle + 0 + 0 + 0$$

$$+ 2\mathrm{Re}\left(bc^* \sum_{i=1}^{n} \langle \phi_n^{\oslash} | \mathbf{S}_i | 0 \cdots 0 \rangle\right) + 0 + |b|^2\Delta + |c|^2\Delta$$

$$+ 2\mathrm{Re}\left(-ab^* \langle \xi_n^{\oslash} | \phi_n^{\oslash} \rangle 2\beta_{n+1}\right) + 0 + 0$$

$$\geq |a|^2 E_n + |b|^2 E_n + 2\mathrm{Re}\left(bc^* \sum_{i=1}^{n} \langle \phi_n^{\oslash} | \mathbf{S}_i | 0 \cdots 0 \rangle\right)$$

$$+ |b|^2\Delta + |c|^2\Delta - 4|a||b|\beta_{\max}, \tag{3.36}$$

where we lower bounded the last term using absolute values, a maximum magnitude of the β's, and $|\langle \psi_n^{\oslash} | \phi_n^{\oslash} \rangle| \leq 1$. Next, we observe that the term $\langle \phi_n^{\oslash} | \mathbf{S}_i | 0 \cdots 0 \rangle_n$ is nonzero only for parts of $|\phi_n^{\oslash}\rangle$ with a single $|1\rangle$. The largest magnitude it could possibly have is when the state $|\phi_n^{\oslash}\rangle$ is made *only* from states with a single $|1\rangle$ as $\frac{1}{\sqrt{n}} \sum_{i=1}^{n} |0 \cdots 1_i \cdots 0\rangle$. We then get $\sum_{i=1}^{n} \langle \phi_n^{\oslash} | \mathbf{S}_i | 0 \cdots 0 \rangle \geq -2\beta_{\max}\sqrt{n}$. Putting this in, recalling (3.33) and using absolute values, we get

$$E_{n+1} \geq \left(|a|^2 + |b|^2\right)\left(\frac{3\Delta}{4} - \delta_n\right) - |b||c| \cdot 4\beta_{\max}\sqrt{n} + \left(|b|^2 + |c|^2\right)\Delta - |a||b| \cdot 4\beta_{\max} \tag{3.37}$$

$$= \frac{3\Delta}{4}\left(|a|^2 + |b|^2 + |c|^2\right) - \left(|a|^2 + |b|^2\right)\delta_n + |b|^2\Delta - |a||b| \cdot 4\beta_{\max}$$

$$- |b||c| \cdot 4\beta_{\max}\sqrt{n} + |c|^2 \cdot \frac{\Delta}{4} \tag{3.38}$$

$$\geq \frac{3\Delta}{4} - \delta_n + |b|^2\Delta - |b| \cdot 4\beta_{\max} + \underbrace{|c|\left(|c|\frac{\Delta}{4} - |b| \cdot 4\beta_{\max}\sqrt{n}\right)}_{f(|c|)}, \tag{3.39}$$

where we have used (3.33), and then $|a|^2 + |b|^2 \leq 1$ and $|a| \leq 1$. Independent of $|b|$, let us look at $f(|c|)$, a quadratic function of $|c|$. Its minimum is at $|c| = \frac{|b| \cdot 8\beta_{\max}\sqrt{n}}{\Delta}$, with the value $f_{\min} = -\frac{|b|^2 \cdot 16n\beta_{\max}^2}{\Delta}$. In (3.39) it means

$$E_{n+1} \geq \frac{3\Delta}{4} - \delta_n + |b|^2\Delta - |b| \cdot 4\beta_{\max} - \frac{16|b|^2 n\beta_{\max}^2}{\Delta} \tag{3.40}$$

$$= \frac{3\Delta}{4} - \delta_n + |b|^2 \underbrace{\left(\Delta - \frac{16n\beta_{\max}^2}{\Delta}\right)}_{\geq 9\Delta/10} - |b| \cdot 4\beta_{\max} \tag{3.41}$$

$$\geq \frac{3\Delta}{4} - \delta_n + \underbrace{|b|\left(|b|\frac{9\Delta}{10} - 4\beta_{\max}\right)}_{g(|b|)}, \tag{3.42}$$

where in the second line we have used $\Delta \geq 160M\gamma_{\max}$ to guarantee $\Delta - \frac{16\beta_{\max}^2 n}{\Delta} \geq \Delta - \frac{16\beta_{\max}^2 N}{\Delta} = \Delta - 16M\gamma_{\max} \geq \frac{9\Delta}{10}$. The expression $g(|b|)$ is quadratic in $|b|$, minimized at $|b| = \frac{20\beta_{\max}}{9\Delta}$, giving the value $g_{\min} = -\frac{40\beta_{\max}^2}{9\Delta}$. Putting it into (3.42), we get

$$E_{n+1} \geq \frac{3\Delta}{4} - \delta_n - \frac{40\beta_{\max}^2}{9\Delta} = \frac{3\Delta}{4} - \delta_{n+1}, \tag{3.43}$$

which proves our induction step, as $\delta_n = \frac{40n\beta_{\max}^2}{9\Delta}$. Therefore, (3.33) holds. Let $n = N$ and we have for any $|\psi\rangle \in \mathcal{L}_2$,

$$\begin{aligned}\langle\psi|\widetilde{\mathbf{H}}_0'|\psi\rangle \geq E_N &\geq \frac{3\Delta}{4} - \frac{40N\beta_{\max}^2}{9\Delta} = \frac{3\Delta}{4} - \frac{40MR\beta_{\max}^2}{9\Delta} \\ &= \frac{3\Delta}{4} - \frac{20M\gamma_{\max}}{9} \geq \left(\frac{3}{4} - \frac{1}{72}\right)\Delta = \frac{53}{72}\Delta > \frac{\Delta}{2} = \lambda_*,\end{aligned} \tag{3.44}$$

where in the last line we have used (3.11) and $\Delta \geq 160M\gamma_{\max}$. Combining the above statement with (3.31), we have $\langle\psi|\widetilde{\mathbf{H}}|\psi\rangle > \lambda_*$ for any $|\psi\rangle \in \mathcal{L}_2$.

This concludes the proof of Claim 1.

3.3.2 The perturbation series converges.

Let us now state and prove our second claim – the convergence of the perturbation series for our gadget construction.

Claim 2 *Consider the 2-body gadget Hamiltonian* $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ *defined in* (3.13) *with spectral gap* Δ *between the ground and the first excited subspace of* $\mathbf{H}$*, and a target Hamiltonian* $\mathbf{H}_{targ} = \mathbf{H}_{else} + \sum_{j=1}^{M} \gamma_j \mathbf{A}_j \otimes \mathbf{B}_j$ *with* $\gamma_j = O(1)$ *and* $\mathbf{H}_{else}$ *positive semi-definite. Choose a constant parameter* $d \in (0,1)$ *and an error tolerance* ϵ*. If we set* $\Delta = M^3 R^d$ *and choose the number of direct ancillas per target term* R *and the core size* C *according to*

$$R = \Omega\left(\max\left\{\epsilon^{-\frac{2}{d}}, \left(\frac{\|\mathbf{H}_{else}\|^2}{2M^4\gamma_{max}}\right)^{\frac{1}{d}}, \left(M^3\epsilon^{-2}\right)^{\frac{1}{1-d}}\right\}\right), \qquad C = \Omega\left(M^3 R^d\, \epsilon^{-1}\right), \tag{3.45}$$

then the strengths of the interaction terms in the gadget Hamiltonian are small, i.e. $\beta_j, J = O(\epsilon)$.

Furthermore, the self energy expansion (3.2) *satisfies*

$$\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{targ} \otimes \mathbf{\Pi}_-\| = O(\epsilon), \tag{3.46}$$

where $\mathbf{\Pi}_-$ *is the projector onto* $\mathcal{L}_-$*, and* z *obeys* $|z| \leq \epsilon + \|\mathbf{H}_{else}\| + \sum_{j=1}^{M} |\gamma_j|$.

This claim is one of the central results of this work – it shows that our gadget Hamiltonian (for a 2-local target Hamiltonian) uses only interactions of strength $O(\epsilon)$, *i.e.* no strong interactions. This is qualitatively different from previous constructions which require interactions of strength poly(ϵ^{-1}). However, the price we pay for avoiding strong interactions is that the number of ancillas scales as poly(ϵ^{-1}), as shown in (3.45), while previous constructions require some number of ancillas independent of ϵ. Hence we present a tradeoff between interaction strength and ancilla number in a gadget Hamiltonian.

Let us prove Claim 2. First we show that $\widetilde{\mathbf{H}}$ consists of only weak interaction terms. When we choose $\Delta = M^3 R^d$ for some $d \in (0, 1)$ and substitute it into (3.11), we find that the interaction strength between the target spins and direct ancillas will be $\beta_j = \sqrt{\frac{\gamma_j \Delta}{2R}} = O(\epsilon)$, if we choose

$$R \gg \left(M^3 \epsilon^{-2}\right)^{\frac{1}{1-d}}. \tag{3.47}$$

Next, recalling $\Delta = CJ$, the strength of the interaction J between the core ancillas will be $O(\epsilon)$ if we choose $C \gg M^3 R^d \epsilon^{-1}$.

Furthermore, once we set $\Delta = M^3 R^d$, we can easily satisfy the requirement $\Delta \geq 160 M \gamma_{\max}$ in Claim 1 for reasonable R – more specifically, we need $R \gg (160\gamma_{\max}/M^2)^{1/d}$.

We will now analyze the higher order terms in the self energy expansion $\mathbf{\Sigma}_-(z)$ according to (3.2) and show that the error term in Eq. 3.2 scales as $O(\epsilon)$. The perturbative expansion of $\mathbf{\Sigma}_-(z)$ for the construction in (3.13) yields

$$\mathbf{\Sigma}_-(z) = \mathbf{H}_{\text{else}} + \frac{1}{z-\Delta}\sum_{j=1}^{M} R\beta_j^2(\mathbf{A}_{a_j} - \mathbf{B}_{b_j})^2 + \underbrace{\sum_{k=1}^{\infty} \mathbf{V}_{-+}\left(\mathbf{G}_+\mathbf{V}_+\right)^k \mathbf{G}_+\mathbf{V}_{+-}}_{\text{error}}. \quad (3.48)$$

We can associate every term in the perturbation series with a *path* starting in the ancilla state $|0\rangle_w|0\rangle_{\mathcal{C}}$ (i.e. belonging to $\mathcal{L}_-$) to states in $\mathcal{L}_+$ and back to $\mathcal{L}_-$. Each path consists of a sequence of virtual *transition steps* between states of the ancillas, denoted $x \to x'$ with R-bit strings x, x'. The number of steps for a path is dependent on the order of the perturbation term. A path for the m-th order is

$$\mathcal{L}_- \xrightarrow{\mathbf{V}_{-+}} |y\rangle \underbrace{\xrightarrow{\mathbf{V}_+} |y_1\rangle \xrightarrow{\mathbf{V}_+} |y_2\rangle \xrightarrow{\mathbf{V}_+} \cdots \xrightarrow{\mathbf{V}_+} |y_{m-2}\rangle \xrightarrow{\mathbf{V}_+}}_{(m-2)\text{ steps}} |y'\rangle \xrightarrow{\mathbf{V}_{+-}} \mathcal{L}_-, \quad (3.49)$$

where y and y' are R-bit strings with Hamming weight 1, and $|y_i\rangle \in \mathcal{L}_+$. In particular, these states belong to the subspace $\mathcal{L}_2 = \mathcal{K}_-^{\perp} \otimes \mathcal{S}_0$ in (3.19). Observe that each term in $\mathbf{T}_m = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{m-2}\mathbf{G}_+\mathbf{V}_{+-}$ is composed from transitions of the following three types

1. a $|0\rangle \to |1\rangle$ flip of some direct ancilla qubit w,
2. a $|1\rangle \to |0\rangle$ flip of some w,
3. the state of the ancillas stays the same.

In the first two cases, $\mathbf{V}_+$ (also $\mathbf{V}_{-+}$ or $\mathbf{V}_{+-}$) contributes the term from $\mathbf{V}$ that flips the direct ancilla w via $\cdots \otimes \mathbf{X}_w$. In the third case, the ancilla state stays the same, and $\mathbf{V}_+$ contributes a term that contains interaction with w via $\cdots \otimes |1\rangle\langle 1|_w$. This type of term contains the factor $\mathbf{H}_{\text{else}}$. Note that for the $k-2$ transitions, the number of flips k_f cannot exceed k. Furthermore, it must be even for the transition to terminate in $\mathcal{L}_-$. Finally, every transition step $y_i \to y_{i+1}$ also contributes a factor $\frac{1}{z-h(y_i)\Delta}$ coming from $\mathbf{G}_+$, with $h(y_i)$ the Hamming weight of the string y_i.

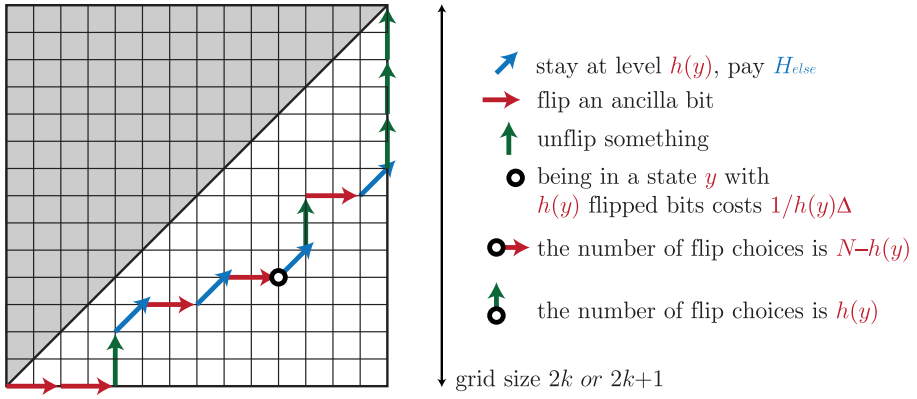


Fig. 3.5. A graphical representation of the contributions to the error term of order $m = 2k$ or $m = 2k + 1$. An up- and right-moving, sub-diagonal path corresponds to a sequence of transitions. A bit flip moves 2 squares horizontally/vertically, while "staying" moves across one square diagonally. The distance from the diagonal corresponds to the number $h(y)$ of flipped ancillas in a given state y.

We can find the norms of the perturbation terms at a given order by enumerating all possible paths and adding up their contributions. For this, we introduce a graphical representation of the paths in Fig. 3.5. Each grid point in the lower-right triangle, including the diagonal points, corresponds to a state with a particular number of the direct ancillas flipped. We start from the lower-leftmost point, which corresponds to the all-zero subspace $\mathcal{L}_-$. Each transition (ancilla flip) maps to a rightwards or upwards movement on the graph, while remaining in the high-energy subspace is depicted by a diagonal step. A valid path ends at the top-rightmost point, which again belongs to the ground state subspace $\mathcal{L}_-$. Furthermore, a valid path can touch the diagonal line only at the last step of the transition (otherwise, it would be a composition of paths at lower orders).

Suppose at a certain point the direct ancillas are in a state $|y\rangle$ with $h(y)$ ancillas in $|1\rangle$ and the rest in $|0\rangle$, with $h(y) \in \{1, 2, \cdots, N\}$ being the Hamming weight of y. Let us first look at transition steps that flip an ancilla takes $|y\rangle$ to a new state $|y'\rangle$ where y and y' differ by one bit.

1. If an ancilla in $|y\rangle$ is flipped from $|0\rangle$ to $|1\rangle$, we move to the right in Figure 3.5. There are $N-h(y)$ ways to flip a 0 to 1 at this point, and we simply overestimate it by N. Furthermore, we get a contribution from $\mathbf{G}_+$, and we overestimate it by $\|\mathbf{G}_+\| \leq \frac{1}{|h(y)\Delta - z|} \leq \frac{1}{\Delta}$. Thus, we find that the norm of a contribution from this first type of transition step is upper bounded by $\frac{N}{\Delta}$.

2. Second, when an ancilla is flipped from $|1\rangle$ to $|0\rangle$, we move up in Figure 3.5. There are $h(y)$ ways to unflip a spin now. The resolvent $\mathbf{G}_+$ again contributes a factor $\frac{1}{z-h(y)\Delta}$. Taken together, the factor $h(y)$ "cancels"[11] the $h(y)$ in the denominator from the resolvent. The contribution from this process is less than $\frac{1}{\Delta}$.

3. Third, for a step that keeps the ancilla state, we remain in the same state y, and move one step diagonally on the graph, getting a contribution $\mathbf{H}_{\text{else}}$. We can

[11] Here "cancel" means that the product $\frac{h(y)}{|h(y)\Delta - z|}$ is $O(\Delta^{-1})$.

do this in $h(y)$ ways, because there are $h(y)$ ancillas in the state $|1\rangle$ that terms like $\cdots \otimes |1\rangle\langle 1|$ apply to. The resolvent $\mathbf{G}_+$ again contributes a factor $\frac{1}{z-h(y)\Delta}$. Similarly to what we did above, we "cancel" the factor $h(y)$, and conclude that a contribution from this type of step is upper bounded by $\frac{\|\mathbf{H}_{\text{else}}\|}{\Delta}$.

Altogether, at order m, our paths have length m, out of which we have f flips, f unflips, and $m-2f$ diagonal steps. The contribution of each such path is upper bounded by

$$c_{\text{path}} \leq \frac{N^f (2\beta_{\max})^{2f} \|\mathbf{H}_{\text{else}}\|^{m-2f}}{\Delta^{m-1}}. \tag{3.50}$$

We now need to find an upper bound on the number of valid paths such as the one shown in Fig. 3.5.

If we did not have the diagonal steps, for even $m = 2k$, this would be the k^{th} Catalan number – the number of up- & right-moving paths between corners of a square of size $2k$ that don't pass above the diagonal. In our case, the situation is just a bit more difficult. The number of $2k$-step (resp. $(2k+1)$-step) paths is upper bounded by the *Motzkin number* of order $2k$ (resp. $2k+1$). These numbers correspond to a number of up-, diagonal-, and right-moving paths across a square, remaining below the diagonal. It suffices for our purposes to use a crude upper bound on the Motzkin numbers: $M_{2k} \leq 3^{2k}$ and $M_{2k+1} \leq 3 \cdot 3^{2k}$, basically saying we have ≤ 3 ways to go at each step. This is grossly over-counting (e.g. going above the diagonal, going farther from the diagonal than N, etc.), but we do not mind, as it will suffice for our argument. Let us finish it first for even $m = 2k$ and then for odd $m = 2k+1$.

Upper bounds on the $(2k)^{th}$ *order.* In estimating the error, we here consider only the 4^{th} order and onward, i.e. $k \geq 2$, as the second order is the actual term that we want to generate (for details of the 2nd order, see Appendix B).

In order to make sure that the sequence of transitions finishes at $\mathcal{L}_-$, the number of flips $k_f = 2f$ must be even ($f \in \mathbb{N}$, $f \leq k$). Hence, the number of steps where the

ancilla state stays the same is $2(k-f)$, an even number. A contribution from some path to $\mathbf{\Sigma}_{-}(z)$ is a term whose norm is upper bounded by

$$\leq \frac{\overbrace{N(2\beta_{\max})^2 \times \cdots \times N(2\beta_{\max})^2}^{f} \times \overbrace{\|\mathbf{H}_{\text{else}}\|^2 \times \cdots \times \|\mathbf{H}_{\text{else}}\|^2}^{k-f}}{\Delta^{2k-1}}. \tag{3.51}$$

The condition $R \geq \left(\frac{\|\mathbf{H}_{\text{else}}\|^2}{2M^4\gamma_{\max}}\right)^{\frac{1}{d}}$ from (3.45), combined with $N = MR$ and (3.11) implies

$$\|\mathbf{H}_{\text{else}}\|^2 \leq 2M^4R^d\gamma_{\max} = 4MR\frac{\Delta\gamma_{\max}}{2R} = N\left(2\beta_{\max}\right)^2. \tag{3.52}$$

Using this and (3.11), we conclude that the overall contribution of a single path (3.51) is bounded from above by

$$\leq \left(\frac{N(2\beta_{\max})^2}{\Delta}\right)^k \frac{1}{\Delta^{k-1}} = 2^k\left(M\gamma_{\max}\right)^k \frac{1}{\Delta^{k-1}} = \Delta\left(\frac{2M\gamma_{\max}}{\Delta}\right)^k. \tag{3.53}$$

The total number of legal paths is less than 9^k. Thus, the norm of the $(2k)^{\text{th}}$ order is bounded from above by

$$\|\mathbf{T}_{2k}\| \leq 9^k\Delta\left(\frac{2M\gamma_{\max}}{\Delta}\right)^k = \Delta\left(\frac{18M\gamma_{\max}}{\Delta}\right)^k. \tag{3.54}$$

We have chosen $\Delta \gg M$, which makes it a small contribution, as we wanted.

Upper bounds on the $(2k+1)^{th}$ *order.* Finding a bound on the 3^{rd} order is straightforward:

$$\|\mathbf{T}_3\| = N\cdot(2\beta_{\max})\cdot\frac{1}{\Delta}\cdot\|\mathbf{H}_{\text{else}}\|\cdot\frac{1}{\Delta}\cdot(2\beta_{\max}) \leq \frac{(4N\beta_{\max}^2)^{\frac{3}{2}}}{\Delta^2} = \sqrt{\frac{(2M\gamma_{\max})^3}{\Delta}}, \tag{3.55}$$

using (3.11). Recalling $\Delta = M^3R^d$, we get $\|\mathbf{T}_3\| \leq (2\gamma_{\max})^{3/2}R^{-d/2} = O\left(R^{-d/2}\right)$, a small contribution.

Analogously, we do the calculation for the general $(2k+1)^{\text{th}}$ order, obtaining

$$\|\mathbf{T}_{2k+1}\| \leq 3\cdot 9^k\cdot\frac{[N(2\beta_{\max})^2]^f\cdot\|\mathbf{H}_{\text{else}}\|^{2(k-f)+1}}{\Delta^{2k}} \leq 3\cdot 9^k\cdot\frac{[N(2\beta_{\max})^2]^{k+\frac{1}{2}}}{\Delta^{2k}}$$

$$\leq 3 \cdot 2^k \cdot 9^k \cdot \left(\frac{M\gamma_{\max}}{\Delta}\right)^k \sqrt{2M\gamma_{\max}\Delta} = \left(3\sqrt{2}\right)\Delta\left(\frac{18M\gamma_{\max}}{\Delta}\right)^{k+\frac{1}{2}}. \tag{3.56}$$

Comparing with (3.55), we find that the last expression is also true for $k = 1$. Therefore, together with (3.54), we can bound all of the terms in the error series by

$$\|\mathbf{T}_m\| \leq 3\sqrt{2}\,\Delta\left(\frac{18M\gamma_{\max}}{\Delta}\right)^{\frac{m}{2}} = 3\sqrt{2}\Delta q^m, \tag{3.57}$$

for $m \geq 3$ with $q = \sqrt{18M\gamma_{\max}/\Delta} = O\left(M^{-1}R^{-\frac{d}{2}}\right)$. Thus, the whole series $\sum_{m=3}^{\infty}\|\mathbf{T}_m\|$ is upper bounded by a geometric series that converges, implying

$$\sum_{m=3}^{\infty}\|\mathbf{T}_m\| \leq const. \times \Delta q^3 = O\left(R^{-\frac{d}{2}}\right) \leq \epsilon, \tag{3.58}$$

for our choice of ϵ when we choose a suitably large $R \gg \epsilon^{-\frac{2}{d}}$. This concludes the proof of Claim 2. □

In conclusion, in Eq. 3.10 we have $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\| = O(\epsilon)$ where the effective Hamiltonian $\mathbf{H}_{\text{eff}} = \mathbf{H}_{\text{targ}} \otimes \mathbf{\Pi}_- + \gamma\mathbf{\Pi}_-$ (up to an overall shift) captures the target Hamiltonian. Therefore we have proven Theorem 3.1.1. Let us have a last look at the required resources:

Remark 3.3.1 *If $\mathbf{H}_{targ}$ acts on n qubits, our gadget Hamiltonian $\widetilde{\mathbf{H}}$ acts on*

$$\begin{aligned} &n + MR + C \\ &\gg n + MR + M^3R^d\,\epsilon^{-1} \\ &\gg n + \max\left\{M\epsilon^{-\frac{2}{d}} + M^3\epsilon^{-3},\ \left(M^{4-d}\epsilon^{-2}\right)^{\frac{1}{1-d}},\ M\left(\frac{\|\mathbf{H}_{else}\|^2}{2M^4\gamma_{max}}\right)^{\frac{1}{d}} + \epsilon^{-1}\,\frac{\|\mathbf{H}_{else}\|^2}{2M\gamma_{max}}\right\} \end{aligned} \tag{3.59}$$

qubits. If the interaction graph of $\mathbf{H}_{targ}$ has degree D, then the interaction graph of the gadget Hamiltonian has total degree $\max\{DR, RC\} = poly(D, \epsilon^{-1}, \|\mathbf{H}_{else}\|, M)$.

This concludes the story of the 2-body gadgets with weak interactions. Let us now apply the construction to reducing k-local to 2-body with weak interaction ($k \geq 3$), and prove Corollary 1.

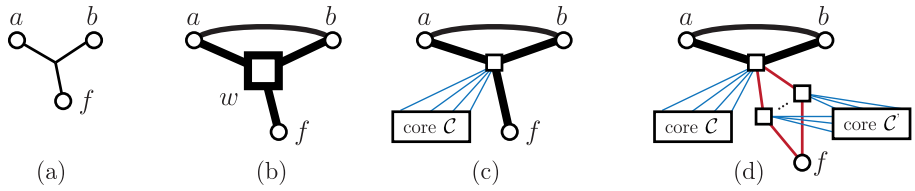


Fig. 3.6. 3-local interactions from weak interactions. (a): The 3-local interaction we want to approximate. (b): The standard construction by Oliveira and Terhal [8] with target term $\mathbf{A}_a \otimes \mathbf{B}_b \otimes \mathbf{F}_f$ replaced by one (direct) ancilla w in a large field Δ, interacting with the target spins via strong interactions of order $\Delta^{2/3}$. In addition, a and b interact with strength of order $\Delta^{1/3}$ to compensate for the error term at 2^{nd} order perturbation theory. (c): The local fields are replaced by interactions with a core. (d): Each strong 2-local interaction term can be reduced to many $O(1)$ terms by our 2-body gadget construction, using another common core.

3.4 Reducing k-body to 2-body interactions ($k \geq 3$)

With the new 2-body construction in mind, is it possible to use the *core* idea and "parallelism" of the 2-body gadgets to construct a 3-body to 2-body gadget that also uses only weak interactions? There is a straightforward way to combine the usual 3-to-2-body gadgets with our strong-from-weak 2-body construction as sketched in Figure 3.6. This is what we claim in Corollary 1. We start from the usual 3-body to 2-body construction in [8] and replace the strong 1-local term of magnitude Δ by interactions with a core. Finally, we reduce the large-norm 2-body interactions in these gadgets with weak ones using the 2-body gadgets from Section 3.3.

For general k-body to 2-body reduction, we can resort to the construction from [9], where the gadget Hamiltonian consists of only 2-local interaction terms (*i.e.* no extra 1-local terms). This makes it easy to directly apply our new 2-body gadgets and reduce the gadget Hamiltonian to one with only weak interactions.

Although it is possible to apply our construction to reduce any k-body target Hamiltonian to a 2-body one with arbitrarily weak interactions, the qubit overhead is

likely exponential in k. With the original constructions proposed in [7,8] it is possible to use $O(k)$ qubits with an exponential overhead in interaction strengths [1]. Perhaps a middle ground between the two constructions could be sought such that both the interaction strength and qubit overheads are polynomial in k.

3.5 Conclusion

A gadget construction based on perturbation theory allows us to map between Hamiltonians of different types, with the same low-lying spectral properties. First, we replace strong interactions by *repetition* of interactions with "classical" ancillas; it works because for a low-energy state, all our extra qubits are close to the state $|0\rangle$. This is reminiscent of repetition encoding found e.g. in [150]. Second, we employ *parallelization*; it is crucial to show that the perturbation series converges even with many gadgets, relaxing the usual assumption about the norm of the perturbation.

This construction should find use in computer science as well as physics. First, in complexity theory, Theorem 3.1.1 together with [7] or Corollary 1 with [125] implies QMA-completeness of the 2-local Hamiltonian problem with non-repeated terms with norm at most $O(1)$ and an $O(1)$ promise gap. As a consequence, we also obtain efficient universality for quantum computation with time-independent, 2-local Hamiltonians with restricted form/strength of terms, complementing [8, 151, 152]. Second, our amplification method from Corollary 2 has been utilized in a counterexample to the generalized area law in [3]. Finally, we envision practical experimental applications of Theorem 3.1.1 – strengthening effective interactions between target (atomic) spins through many (but even for a few R) coupled mediator spins. In our case, these interactions need to be precisely tuned, while elsewhere we have seen disordered networks used to enhance transport between two sites in a quantum system [153].

Thinking further about interaction strengths and spectra of local Hamiltonians, we realize that Corollary 2 allows us to amplify the *eigenvalue gap* (low eigenvalue spacing) of a Hamiltonian. Does it have direct implications for hardness of Local

Hamiltonian problems? When we use it on Hamiltonians appearing in QMA-complete constructions, the fractional *promise gap* (the ratio of the number of frustrated terms to the number of all terms in the Hamiltonian for a ground state of a local Hamiltonian) gets smaller. Thus, it does not directly help us move towards the quantum PCP conjecture [143]. Nevertheless, we have added another tool for mapping between Hamiltonians to our repertoire.

An important problem remains open. The price we pay for our construction is a massive blowup in the degree (the number of interactions per particle). Is there a possibility of a quantum degree-reduction gadget? One might try to use a "bad" quantum code for encoding each spin into several particles, whose encoded low-weight operators that can be implemented in *many* possible ways; this does not seem possible for both X and Z operators. As things stand, without a degree-reduction gadget, we do not have a way to reproduce our results in simpler geometry. It would be really interesting if one indeed could create $O(1)$-norm effective interactions from $O(1)$-terms in 3D or even 2D lattices.

We also need to think about the robustness of our results – what will change when the Hamiltonians are not exactly what we asked for? How precise do we need to be (e.g. for the 3-body to 2-body gadgets), so that the second- and first-order terms get canceled? Also, Bravyi, Terhal, DiVincenzo and Loss [129] mentioned that a k-body to 2-body reduction might possibly be implemented with poly(k) overhead in interaction strength (instead of exponential in k). However, this question remains open. The exponential scaling in the overhead in [129] is due to the usual gadget constructions which require poly(ϵ^{-1}) interaction strength. We hope (but haven't proven) that with our new gadget construction, this result could be improved.

4. EFFICIENT ALGORITHMS FOR ESTIMATING PERTURBATIVE ERROR

We introduced perturbation theory in Section 1.3. Also we have showed how the amount of information needed for specifying the state of a many-body system in quantum mechanics commonly scales exponentially as the system size (Section 1.2). This poses a fundamental difficulty in using perturbation theory at arbitrary order. As one computes the terms in the perturbation series at increasingly higher orders, it is often important to determine whether the series converges and if so, what is an accurate estimation of the total error that comes from the next order of perturbation up to infinity. Here we present a set of efficient algorithms that compute tight upper bounds to perturbation terms at arbitrary order. We argue that these tight bounds often take the form of symmetric polynomials on the parameter of the quantum system. We then use cellular automata as our basic model of computation to compute the symmetric polynomials that account for all of the virtual transitions at any given order. At any fixed order, the computational cost of our algorithm scales *polynomially* as a function of the system size. We present a non-trivial example which shows that our error estimation is nearly tight with respect to exact calculation.

4.1 Overview

An overwhelming majority of problems in quantum physics and quantum chemistry do not admit exact, analytical solutions. Therefore one has to resort to approximation methods based on for instance series expansions [39, 154–157]. Often these expansions are truncated to a finite order r as an approximation of the true solution an the remaining terms from the $(r+1)^{\text{th}}$ order on are errors. It is then important to estimate the magnitude of errors at arbitrary order as a gauge of how

the series performs as an approximate solution. The main challenge in this task is that exact calculation of the perturbative terms commonly scales exponentially as the size of the system under consideration, making it hard to pinpoint the regime where perturbation theory yields acceptable accuracy [39].

Here we present an efficient method for deriving tight upper bounds for the norm of perturbative expansion terms at arbitrary order. As we have mentioned in Section 1.3, the use of perturbation theory starts with identifying a physical system $\widetilde{\mathbf{H}}$ as a sum of an unperturbed Hamiltonian $\mathbf{H}$ that acts on a Hilbert space $\mathcal{H}$ and a perturbation $\mathbf{V}$. As shown in Figure 4.1a, we assume that $\mathbf{H} = \mathbf{H}^{(1)} + \mathbf{H}^{(2)} + \cdots + \mathbf{H}^{(m)}$ consists of m identical and non-interacting unperturbed subsystems with Hilbert space $\mathcal{H}^{(i)}$, $i = 1, \cdots, m$. Each subsystem interacts with a "bath" $\mathcal{B}$ through perturbation $\mathbf{V}$ that is presumably small. We further assume that for each subsystem $\mathbf{H}^{(i)}$, $\mathbf{V}$ can only cause transitions in neighboring energy levels (Figure 4.1b). This form of physical setting is typical in for example spin systems with perturbation on individual spins via local fields [158, 159], or in Hartree approximation where m identical particles interact with a mean field [155]. Here $\mathbf{V}$ does not necessarily act identically on each $\mathcal{H}^{(i)} \otimes \mathcal{B}$ for every i. For a given $\mathbf{V}$, one could determine an upper bound λ_i for each subsystem i such that $|\langle\phi|\mathbf{V}|\phi'\rangle| \leq \lambda_i$ for any $|\phi\rangle$, $|\phi'\rangle$ being eigenstates of $\mathbf{H}^{(i)}$. We could also determine an upper bound ω such that for any $|\phi\rangle$ that is an eigenstate of $\mathbf{H}$, $|\langle\phi|\mathbf{V}|\phi\rangle| \leq \omega$. With the spectrum of each $\mathbf{H}^{(i)}$ fully known, one could also determine for each energy level s and t the maximum number of possible ways for an eigenstate at energy level s to make a transition to a state of energy level t via the perturbation $\mathbf{V}$. We let this number be M_{st} for all $\mathbf{H}^{(i)}$, since their spectra are identical.

In many cases we are only concerned about the property of the effective Hamiltonian below certain cutoff energy E_*. Similar to Section 1.3, we assume that the ground state energy of every $\mathbf{H}^{(i)}$ is 0 and $E_* = \Delta/2$ where $\Delta = E_1$ is the spectral gap between the ground and the first excited state. For $\|\mathbf{V}\|$ small enough compared to Δ we could extract this information using the operator valued resolvent

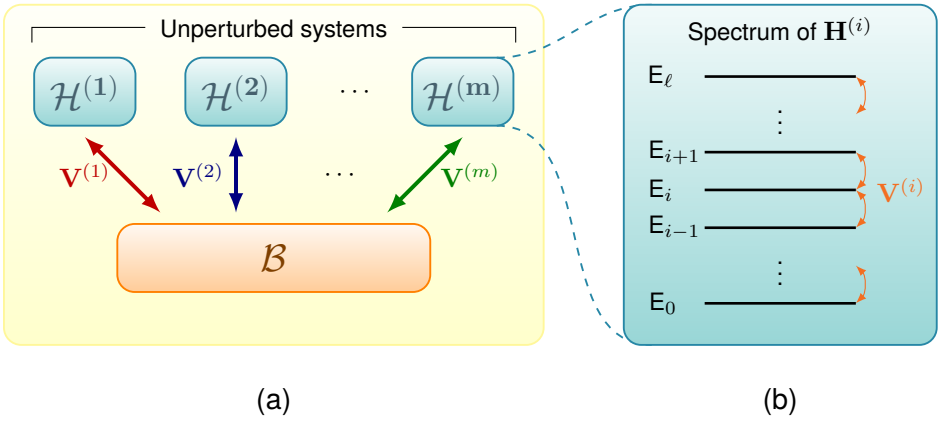


Fig. 4.1. General setting of the perturbation theory.

$\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$ with a small expansion parameter z and construct the self-energy (we duplicate Equation 1.26 here)

$$\mathbf{\Sigma}_-(z) = \mathbf{H}_{--} + \mathbf{V}_{--} + \mathbf{V}_{-+}\mathbf{G}_{++}\mathbf{V}_{+-} + \mathbf{V}_{-+}\mathbf{G}_{++}\mathbf{V}_{++}\mathbf{G}_{++}\mathbf{V}_{+-} + \cdots \tag{4.1}$$

where we partition $\mathcal{H}$ into subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$, with $\mathcal{L}_-$ being the subspace of $\mathcal{H}$ spanned by $\mathbf{H}$ eigenstates with energy below E_* and $\mathcal{L}_+$ being the complement of $\mathcal{L}_-$ in $\mathcal{H}$, and let $\mathbf{O}_{\pm\pm} = \mathbf{\Pi}_\pm \mathbf{O} \mathbf{\Pi}_\pm$ be projections of any operator $\mathbf{O}$ onto the $\mathcal{L}_\pm$ subspaces. $\mathbf{\Pi}_-$ and $\mathbf{\Pi}_+$ are projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively. To compute an approximation to the low-energy effective Hamiltonian of $\widetilde{\mathbf{H}}$, one simply truncates Equation 4.1 at low orders to obtain an effective Hamiltonian $\mathbf{H}_{\text{eff}}$ and discard the remaining terms which constitutes the error of the perturbation series. Here we are only restricted to convergent series. For divergent series one may resort to resummation techniques such as Padé approximation [154]. If we denote the r^{th} order term in the self energy expansion (4.1) as $\mathbf{T}_r = \mathbf{V}_{-+}(\mathbf{G}_{++}\mathbf{V}_{++})^{r-2}\mathbf{G}_{++}\mathbf{V}_{+-}$ for $r \geq 2$ and $\mathbf{T}_1 = \mathbf{V}_{--}$, then our effective Hamiltonian $\mathbf{H}_{\text{eff}} = \mathbf{T}_1 + \mathbf{T}_2 + \cdots + \mathbf{T}_R$ for some R and the remaining terms $\mathbf{T}_{R+1} + \mathbf{T}_{R+2} + \cdots$ are error. The connection between the magnitude of the error $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2$ and the spectral difference between $\widetilde{\mathbf{H}}$ and $\mathbf{H}_{\text{eff}}$ is well established. If for a suitable range of z, $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2$ is no greater than ϵ, then the energies of $\mathbf{H}_{\text{eff}}$ are at most ϵ apart from their counterparts in the low energy spectrum of $\widetilde{\mathbf{H}}$ (see [7,8]). Our goal is precisely to find tight upper bounds for the magnitude of the error terms $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2$.

For convergent series it suffices to be able to find tight estimates for the ∞-norm of the r^{th} order term $\|\mathbf{T}_r\|_\infty$ for any $r \geq 2$. The ∞-norm of a matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ is defined as $\max_{i=1,\cdots,m} \sum_{j=1}^{n} |a_{ij}|$. We could bound $\|\mathbf{T}_r\|_\infty$ from above by a function of λ_i, M_{st} and ω. Because $\mathbf{T}_r$ is essentially a matrix product, one could think of the matrix element $\langle\phi|\mathbf{T}_r|\phi'\rangle$ as a sum of r-step walks on the eigenstates of $\mathbf{H}$, which can be written as $|\phi\rangle \to |\phi^{(1)}\rangle \to \cdots \to |\phi^{(r-1)}\rangle \to |\phi'\rangle$, with each $|\phi^{(i)}\rangle$ being an eigenstate of $\mathbf{H}$ and each step of the walk contributing a factor and the total weight

of the walk is the product of all the factors. Using the scalar quantities λ_i, M_{st} and ω symbols we could derive an upper bound to $|\langle\phi|\mathbf{T}_r|\phi'\rangle|$ by noting that

$$\begin{aligned} |\langle\phi|\mathbf{T}_r|\phi'\rangle| \;\leq\; & \sum_{\{|\phi^{(i)}\rangle\}} |\langle\phi|\mathbf{V}|\phi^{(1)}\rangle| \cdot |\langle\phi^{(1)}|\mathbf{G}|\phi^{(1)}\rangle| \cdot |\langle\phi^{(1)}|\mathbf{V}|\phi^{(2)}\rangle| \cdots \\ & \cdots |\langle\phi^{(r-2)}|\mathbf{V}|\phi^{(r-2)}\rangle| \cdot |\langle\phi^{(r-1)}|\mathbf{G}|\phi^{(r-1)}\rangle| \cdot |\langle\phi^{(r-1)}|\mathbf{V}|\phi'\rangle| \end{aligned} \tag{4.2}$$

where the summation is over all possible r-step walks on the eigenstates of $\mathbf{H}$ that starts at $|\phi\rangle$ and ends at $|\phi'\rangle$. The factors $|\langle\phi^{(i)}|\mathbf{G}|\phi^{(i)}\rangle| = 1/|z - E^{(i)}|$, where $E^{(i)} = \langle\phi^{(i)}|\mathbf{H}|\phi^{(i)}\rangle$, can be computed easily since the spectrum of $\mathbf{H}$ is known. Suppose $\mathbf{V}$ transforms an $\mathbf{H}$ eigenstate $|\phi^{(i)}\rangle$ into $\mathbf{V}|\phi^{(i)}\rangle = |\phi^{(i+1)}\rangle$ by changing the energy level of one of the subsystems (say $\mathbf{H}^{(i)}$) from s to t. Then $|\langle\phi^{(i)}|\mathbf{V}|\phi^{(i+1)}\rangle| \leq \lambda_i M_{st}$. However, if $|\phi^{(i)}\rangle = |\phi^{(i+1)}\rangle$, then we have $|\langle\phi^{(i)}|\mathbf{V}|\phi^{(i+1)}\rangle| \leq \omega$. For each walk on the eigenstates of $\mathbf{H}$ we could then assemble an upper bound that looks like for example (Figure 4.2 top layer)

$$\lambda_i M_{st} \cdot \frac{1}{|z - E^{(1)}|} \cdot \lambda_j M_{pq} \cdot \frac{1}{|z - E^{(2)}|} \cdot \omega \cdots . \tag{4.3}$$

At the second order we could use this technique to bound $\|\mathbf{T}_2\|_\infty$ from above as

$$\begin{aligned} \|\mathbf{T}_2\|_\infty \;\leq\; & \lambda_1 M_{01} \cdot \frac{1}{|z - E_1|} \cdot \lambda_1 M_{10} + \lambda_2 M_{01} \cdot \frac{1}{|z - E_1|} \cdot \lambda_2 M_{10} + \cdots \\ & \cdots + \lambda_m M_{01} \cdot \frac{1}{|z - E_1|} \cdot \lambda_m M_{10}. \end{aligned} \tag{4.4}$$

where we recall that E_1 is the first excited state energy of any subsystem $\mathbf{H}^{(i)}$ (Figure 4.1b). Each term in Equation 4.4 with λ_j corresponds to a 2-step walk where the j^{th} subsystem is excited from the ground state (0^{th} energy level) into the first excited state and then transitions back to the ground state energy subspace.

The expressions for the upper bounds to $\|\mathbf{T}_r\|_\infty$ such as on the right hand side of Equation 4.4 looks simple for $r = 2$. At higher order, however, the situation quickly becomes more complicated. Intuitively this is because each unperturbed system has ℓ possible energy levels, and m such subsystems could manifest ℓ^m possible ways in

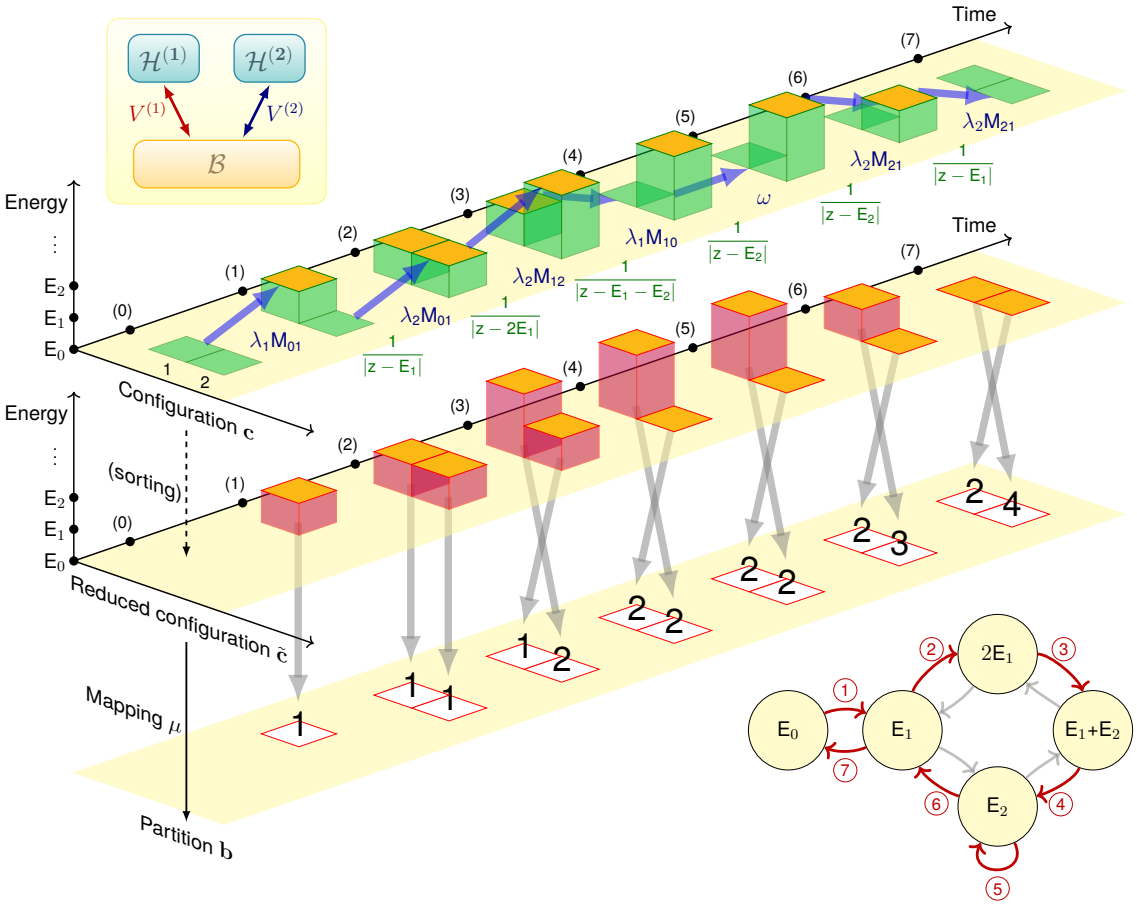


Fig. 4.2. An example of a walk arising at 7^{th} order perturbation theory $\mathbf{T}_7 = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^5\mathbf{G}_+\mathbf{V}_{+-}$. Top left: the specific physical setting concerned, where the number of subsystems is $m = 2$. Top layer: the relationship between the 7-step walk in the space of energy configurations $\mathbf{c}$ and an upper bound associated with it. Each transition due to $\mathbf{V}$ is associated with a factor of either $\lambda_i M_{st}$ or ω. Each intermediate step with energy $E^{(i)}$ contributes a term $1/|z - E^{(i)}|$ due to $\mathbf{G}_+$. Middle layer: the corresponding walk in $\tilde{\mathbf{c}}$, where at each step $\tilde{\mathbf{c}}^{(i)}$ is obtained by sorting $\mathbf{c}$ in descending order. Bottom layer: the corresponding change in the partition $\mathbf{b}$ and the mapping $\mu : \tilde{\mathbf{c}} \mapsto \mathbf{b}$ maintained throughout. By convention, the partition $\mathbf{b}$ is always of non-decreasing order. Bottom right: the walk in the space of energy combination $\mathbf{n}$ corresponding to the walk in $\tilde{\mathbf{c}}$. This walk in $\mathbf{n}$ is what the cellular automaton algorithm essentially implements.

which the energies of each subsystems are assigned. Therefore any matrix element of $\mathbf{T}_r$ should be a sum of roughly at most $O(\ell^{mr})$ walks, yielding an exponential complexity with respect to the total system size m. However, we note that such exponential complexity could be reduced to merely poly(m) by exploiting the inherent permutation symmetry of upper bounds such as Equation 4.4. The essential observation is that these upper bounds are invariant with respect to permutation of the subsystems. This implies that they are *symmetric functions* over the λ_i variables. In particular, these upper bounds to $\|\mathbf{T}_r\|_\infty$ are linear combinations of *monomial symmetric polynomials*, which can be written in form of [160]

$$m_{\mathbf{b}}(\boldsymbol{\lambda}) = \sum_{\pi \in S_k} \lambda_{\pi(1)}^{b_1} \lambda_{\pi(2)}^{b_2} \cdots \lambda_{\pi(k)}^{b_k}$$

where $\mathbf{b} \in \mathbb{N}^{\mathbf{k}}$ is a vector which we call *partition*, $\boldsymbol{\lambda} = (\lambda_1, \cdots, \lambda_m)$ and the summation is over a permutation group S_k, where any permutation π chooses k elements from m elements and permutes them. For example, $m_{(1,2)}(\lambda_1, \lambda_2, \lambda_3) = \lambda_1\lambda_2^2 + \lambda_1\lambda_3^2 + \lambda_2\lambda_3^2 + \lambda_2\lambda_1^2 + \lambda_3\lambda_2^2 + \lambda_3\lambda_1^2$ is a monomial symmetric polynomial. Equation 4.4 could be compactly represented as $\|\mathbf{T}_2\|_\infty \leq \frac{1}{|z-E_1|} M_{01}M_{10}m_{(2)}$. At 4^{th} order we could show that

$$\|\mathbf{T}_4\|_\infty \leq \frac{M_{01}M_{10}\omega^2 m_{(2)}}{|(z-E_1)^3|} + \frac{2M_{01}^2M_{10}^2 m_{(2,2)}}{|(z-E_1)^2(z-2E_1)|} + \frac{M_{01}M_{12}M_{21}M_{10}m_{(4)}}{|(z-E_1)^2(z-E_2)|}. \tag{4.5}$$

By respecting the matrix product structure of $\mathbf{T}_r$, the symmetric polynomial upper bounds such as those in Equations 4.4 and 4.5 turn out to be a much more accurate estimation of the true magnitude of $\|\mathbf{T}_r\|_\infty$ than crude bounds using geometric series such as $\|\mathbf{T}_r\|_2 \leq \|\mathbf{V}\|_2 \cdot \|\mathbf{G}_{++}\|_2 \cdot \|\mathbf{V}\|_2 \cdots \|\mathbf{G}_{++}\|_2 \cdot \|\mathbf{V}\|_2$. In later discussions we will demonstrate this point using numerical examples.

The question then becomes how we may assemble expressions such as (4.4) and (4.5) in an algorithmic fashion. We accomplish this efficiently by using *cellular automata* as the basic data structure. In a nutshell, a cellular automaton is a computational model consisting of a network of basic units called *cells* that are connected by directed edges. Each cell stores some data which represent its current *state*. All

the cells are assigned an initial state and the computation proceeds by evolving each cell using an identical rule for updating its state. The new state of each cell is only dependent on the previous states of the same cell and its neighbors. The study of cellular automata dates back to the 1940s [161], followed by interesting constructions [162–164] and formal, systematic study over the past decades [165,166]. Though computationally rich, the structure of cellular automata considered in these contexts are commonly rather simple, with cells that have discrete sets of possible states and are connected by simple network geometries (such as a 2D grid). In our case, as we will discuss later, the cells in cellular automata store more complex data structures and are connected with often non-planar network geometries. The update rules designed specifically so that the coordination of cells as a whole computes the symmetric polynomial upper bound for $\|\mathbf{T}_r\|_\infty$.

The connection between cellular automata and perturbation theory seems unusual at first glance. However, the connection between cellular automata and random walks is well documented [167–169]. Such connection, combined with our earlier discussion on how the symmetric polynomial upper bounds could arise from summing over walks on the set of $\mathbf{H}$ eigenstates, suggests that one may also be able to use cellular automata for the summation over these walks. One could further think of our task of computing a symmetric polynomial upper bound to $\|\mathbf{T}_r\|_\infty$ as summing over walks in a space of *energy configurations* $\mathbf{c}$, which are m-dimensional vectors of indices ranging from 0 to $\ell - 1$ indicating the energy level of each subsystem in a particular $\mathbf{H}$ eigenstate. In other words, $\mathbf{c} = (c_1, \cdots, c_m) \in \{0, 1, \cdots, \ell - 1\}^m$ and $\langle\phi|H^{(i)}|\phi\rangle = E_{c_i}$ for any particular $\mathbf{H}$ eigenstate $|\phi\rangle$. Therefore each r-step walk in the space of $\mathbf{H}$ eigenstates corresponds to a walk in the space of energy configuration $\mathbf{c}$, which is of size $O(\ell^m)$. We could reduce the size of this space by taking every energy configuration $\mathbf{c}$ and sort its elements to produce a new vector $\tilde{\mathbf{c}}$, which we call *reduced energy configuration*. Like the number of energy levels in $\mathbf{H}$, the set of $\tilde{\mathbf{c}}$ is also of size $O(m^\ell)$, which is polynomial is m assuming ℓ is a constant and intensive property of each subsystem (for instance a spin-1/2 particle has $\ell = 2$ if we are only

concerned with the spin degree of freedom). Each energy level of $\mathbf{H}$ is a sum of the energies of the subsystems: $\langle\phi|\mathbf{H}|\phi\rangle = \sum_{i=0}^{\ell-1} n_i E_i = E(\mathbf{n})$ where E_i is one of the ℓ possible energy levels of a subsystem. We could write each energy level of $\mathbf{H}$ as an ℓ-dimension vector $\mathbf{n} = (n_0, n_1, \cdots, n_{\ell-1})$ which we call *energy combination* (Figure 4.2 middle layer).

With the discussion so far we have reduced the problem of summing over walks on the set of $\mathbf{H}$ eigenstates, whose number scales exponentially with respect to system size parameter m, to one that concerns only with walks on the set of $\mathbf{n}$, which is of only polynomial size in m. In accomplishing this reduction, we introduced the notion of energy configuration $\mathbf{c}$ and reduced energy configuration $\tilde{\mathbf{c}}$. Going from walks in $\mathbf{c}$ to $\tilde{\mathbf{c}}$ is a major step that takes advantage of the permutation symmetry with respect to the m subsystems in the r^{th} order from $\mathbf{T}_r$. We capture this symmetry with the use of symmetric polynomials $m_{\mathbf{b}}(\boldsymbol{\lambda})$. We illustrate this concept in Figure 4.2. We note that the partition $\mathbf{b}$ does not contain all of the information associated with a walk in $\tilde{\mathbf{c}}$. Consider a particular walk on the set of $\mathbf{H}$ eigenstates and its associated weight whose functional form is shown in Equation 4.3, $\mathbf{b}$ only records the number of times that some subsystem is acted on by $\mathbf{V}$, without the information about the order and the energies of the subsystem before and after the action (Figure 4.2 bottom layer). For example the partition $(1, 2)$ means "one of the subsystems is acted on by $\mathbf{V}$ once and another is acted on by $\mathbf{V}$ twice". The expression $m_{(1,2)}(\boldsymbol{\lambda})$ sums over the weights of walks that fits that description. But there are more than one possible walks, be it on the set of $\mathbf{H}$ eigenstates or $\mathbf{c}$ or $\tilde{\mathbf{c}}$, that fits the description. Therefore in order for a symmetric polynomial to accurately represent an upper bound to the contributions to $\langle\phi|\mathbf{T}_r|\phi'\rangle$ from all walks in $\tilde{\mathbf{c}}$, a mapping must be maintained between $\mathbf{b}$ and $\tilde{\mathbf{c}}$ to indicate which subsystem is being acted on at the current step. Figure 4.2 shows an example that illustrates the connection between $\tilde{\mathbf{c}}$, $\mathbf{b}$, and μ to a walk in the configuration space $\mathbf{c}$.

In our construction cellular automata that executes the summation over walks in $\tilde{\mathbf{c}}$, each cell corresponds to an energy level of $\mathbf{H}$. Hence there are in total $O(m^\ell)$

cells. We use the energy combinations $\mathbf{n}$ to uniquely label each cell. Then the cells are connected with directed edges such that cell $\mathbf{n}$ will only be connected to cell $\mathbf{n}'$ if there are eigenstates $|\phi\rangle$, $|\phi'\rangle$ of $\mathbf{H}$ with energy combinations $\mathbf{n}$ and $\mathbf{n}'$ respectively such that $|\langle\phi|\mathbf{V}|\phi'\rangle| \neq 0$. In our algorithm each monomial symmetric polynomial $\xi m_{\mathbf{b}}(\boldsymbol{\lambda})$ is represented with a 4-tuple $(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu)$ where ξ is a scalar quantity indicating the weight of $m_{\mathbf{b}}(\boldsymbol{\lambda})$ in the overall symmetric polynomial upper bound. $\tilde{\mathbf{c}}$ and $\mathbf{b}$ are respectively the reduced energy configuration and partition at the current step of the walk. $\mu : \tilde{\mathbf{c}} \mapsto \mathbf{b}$ is a bijective mapping between $\tilde{\mathbf{c}}$ and $\mathbf{b}$, as justified in previous discussion.

Each cell of the automaton stores a list of 4-tuples $(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu)$ as its state. As shown in Figure 4.4, at each iteration the state of each cell is updated in a two-phase process. In phase I (Figure 4.4a), the list of 4-tuples stored in $\mathcal{S}_{\mathbf{n}}$ is first merged with thosed stored in all of the incident edges to $\mathcal{S}_{\mathbf{n}}$ and then the coefficients of all the 4-tuples in $\mathcal{S}_{\mathbf{n}}$ are multiplied by a factor $1/|z - E(\mathbf{n})|$. The intuition is that each 4-tuple corresponds to a particular walk such as the one shown in Figure 4.2. The multiplication by $1/|z-E(\mathbf{n})|$ essentially accounts for the contribution from $\mathbf{G}_+$ in $\mathbf{T}_r$. In phase II, we account for the contribution from $\mathbf{V}$ terms in $\mathbf{T}_r$ by first computing new 4-tuples with $\tilde{\mathbf{c}}$ that can be generated from the current 4-tuples in $\mathcal{S}_n$ with one application of $\mathbf{V}$, and then distributing the new 4-tuples among the outgoing edges $\mathcal{S}_{\mathbf{n},\mathbf{n}''}$, as shown in Figure 4.4b.

As the cells evolve, the 4-tuples are updated and passed along between the cells so that at the end of r iterations, we could glean the symmetric polynomial upper bound from the states of the cells. The update rules for each cell are designed to maintain the property that at any iteration, each cell $\mathbf{n}$ contains a list of 4-tuples $(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu)$ each of which corresponds to the set of all walks in $\tilde{\mathbf{c}}$ that leads up to a state with energy combination $\mathbf{n}$, and $\xi m_{\mathbf{b}}(\boldsymbol{\lambda})$ is an upper bound to the total contribution of the walks on the set of $\mathbf{H}$ eigenstates that share the same corresponding walk in $\tilde{\mathbf{c}}$. In other words, $\xi m_{\mathbf{b}}(\boldsymbol{\lambda})$ is a sum of expressions such as Equation 4.3 for these walks on the set of $\mathbf{H}$ eigenstates. We are able to rigorously show that with suitable initialization,

after r iterations the cellular automaton is indeed able to find a symmetric polynomial upper bound for $\|\mathbf{T}_r\|_\infty$ similar to that of $\|\mathbf{T}_4\|_\infty$ in Equation 4.5.

We stress that the overall time complexity of our algorithm scales *polynomially* as the system size grows. The degree of the polynomial, however, depends on the order of perturbation theory. For convergent series, the exponential dependence on the order r of perturbation theory could be handled in practice by for instance setting a threshold η such that when the symmetric polynomial upper bound computed by the cellular automaton is below η at some order r_c of perturbation, we bound the remaining terms up to infinity by a geometric series. For different problems and choices of η, the value of r_c may vary. But the overall polynomial scaling with respect to the system size m should not be affected.

In the mathematical developments of physical theories one is often concerned with the *representation* of the solution to a problem. For very few problems are we able to find a close-form, explicit formula as a representation of the solution. Series expansions are then introduced to largely enhance our ability to solve difficult problems far beyond analytical solution, as they allow for representation of a much wider class of mathematical objects. If we think of these representations as efficient procedures that allow us to construct our solution, then in greater generality we could argue that the outputs of efficient algorithms are also valid representations of our solution. Our scheme based on cellular automata essentially produces this type of representation: the symmetric polynomial upper bound to $\|\mathbf{T}_r\|_\infty$ that we have devised is most conveniently expressed in form of an algorithmic output, rather its explicit self as a sum of monomials. A similar example to this situation is perhaps the development of tensor networks as representations for quantum ground states [170–172]. As is the case for our algorithmic development, tensor networks are also intended to cope with the exponential size of Hilbert space as the physical system grows. Using innovative data structures based on tensors, one obtains a polynomial size approximation to the true ground state. The resulting ground state is then most conveniently represented in form of a tensor network rather than its exponential-

size self as a linear combination of basis states. Our cellular automaton algorithm could also be thought of as producing an approximation to $\|\mathbf{T}_r\|_\infty$, in the sense that we replace the action of $\mathbf{V}$ on the unperturbed eigenstates $|\phi\rangle$, $|\phi'\rangle$ of each subsystem i by scalar quantities λ_i and ω, and we use the integers M_{st} to obtain a sketch of the structure of $\mathbf{V}$. Such approximations may seem crude at first sight, but they preserve the combinatorial structure of $\mathbf{T}_r$ as a matrix product, and allow for compact description using symmetric polynomials. We use iteration of cell evolution as a natural means to compute these symmetric polynomials. As a result, the output of our cellular automaton algorithm is the most natural representation for the upper bound to $\|\mathbf{T}_r\|_\infty$ that we have devised.

One of the areas where our algorithm could find direct application is quantum computation. Though perturbation theory has been pervasively used for calculating properties of quantum systems, the lack of efficient and effective methods for estimating the error even for convergent series has cast a wide shadow of uncertainty on these calculations. Such problem becomes ever more imminent when one tries to engineer quantum systems that are intended to meet specific application requirements such as quantum computing [1, 2, 173]. As the implementations of quantum devices scale up and perturbation theory finds its inevitable use in analyzing these devices, it is imperative to have a scalable method for estimating the error in the perturbative expansion.

For example, in quantum simulation one often wishes to construct a two-body physical system $\widetilde{\mathbf{H}}$ whose low energy effective interactions $\mathbf{H}_{\text{eff}}$ are many-body [7–9]. The most general construction of $\widetilde{\mathbf{H}}$ to date that could generate arbitrary many-body dynamics in $\mathbf{H}_{\text{eff}}$ is based on perturbation theory. Here in Figure 4.3 we show one example of such construction with $\mathbf{H}_{\text{eff}} = \alpha_1 \mathbf{X}_1\mathbf{X}_2\mathbf{X}_3 + \alpha_2 \mathbf{X}_2\mathbf{Y}_4\mathbf{Z}_5$ being three-body while $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ is entirely two-body [9]:

$$\mathbf{H} = \mathbf{H}^{(1)} + \mathbf{H}^{(2)}, \qquad \mathbf{H}^{(1)} = \frac{\Delta}{4}(\mathbf{Z}_{u_1}\mathbf{Z}_{u_2} + \mathbf{Z}_{u_2}\mathbf{Z}_{u_3} + \mathbf{Z}_{u_1}\mathbf{Z}_{u_3})$$

$$\mathbf{H}^{(2)} = \frac{\Delta}{4}(\mathbf{Z}_{v_1}\mathbf{Z}_{v_2} + \mathbf{Z}_{v_2}\mathbf{Z}_{v_3} + \mathbf{Z}_{v_1}\mathbf{Z}_{v_3})$$

$$\mathbf{V} = \mathbf{V}^{(1)} + \mathbf{V}^{(2)}, \qquad \mathbf{V}^{(1)} = \mu_1(\mathbf{X}_1\mathbf{X}_{u_1} + \mathbf{X}_2\mathbf{X}_{u_2} + \mathbf{X}_3\mathbf{X}_{u_3})$$
$$\mathbf{V}^{(2)} = \mu_2(\mathbf{Y}_4\mathbf{X}_{v_1} + \mathbf{X}_2\mathbf{X}_{v_2} + \mathbf{Z}_5\mathbf{X}_{v_3}) \quad (4.6)$$

where spins with u_i and v_i labels belong to the two unperturbed subsystems. Here we let Δ be orders of magnitude larger than μ_1 and μ_2 and keep the coefficients μ_1 and μ_2 as $\mu_1 = (\alpha_1\Delta^2/6)^{1/3}, \mu_2 = (\alpha_2\Delta^2/6)^{1/3}$. Perturbative calculation on $\widetilde{\mathbf{H}}$ show that the leading three orders $\mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 = \mathbf{H}_{\text{eff}} \otimes \mathbf{\Pi}$ for some projector $\mathbf{\Pi}$ acting on a Hilbert space separate from that of $\mathbf{H}_{\text{eff}}$. The simulator Hamiltonian $\widetilde{\mathbf{H}}$ is constructed such that the perturbative series converges. In our example $\widetilde{\mathbf{H}}$ consists of only two-body spin interactions and parameters $\omega = 0$, $\lambda_1 = \mu_1$, $\lambda_2 = \mu_2$ and M_{st} can be computed from Figure 4.3d. The cellular automaton in this case is set up as in Figure 4.4. We then proceed to evolve the cellular automaton, gathering outputs from the cells corresponding to the low energy subspace. As shown in Figure 4.5, even with the convergence, simple geometric series upper bounds fail to capture the true magnitude of $\|\mathbf{T}_r\|_\infty$ while the output of our cellular automaton algorithm is essentially tight with respect to the true value. Note that the true value takes an exponential amount of computational effort in m while our cellular automaton algorithm costs only polynomial in m, as discussed before. This implies that we could obtain efficient and accurate estimations for the error of our quantum simulation that are not previously available.

Beyond quantum computing, our algorithm should retain its effectiveness for general spin systems and find its application in greater areas of condensed matter physics. For example, dimensional scaling method, pioneered by Herschbach [157], uses the inverse space dimensionality as a perturbation free parameter to solve complex many-body problems by taking the large-dimensional limit as the zeroth order approximation. At this limit many problems admit a simple solution, as in the electronic structure calculations of atoms and molecules. Moreover, the second-order term also can be calculated but the higher order terms are cumbersome and hard to estimate [157]. This new proposed algorithm might be useful to estimate the perturbation error in

dimensional scaling method which will lead to a very powerful and efficient approach to solve complex many-body problems. Like tensor networks, which triggered an entirely new direction of research, it would be exciting to see what deeper truths of our quantum world could be unveiled by innovative proposals of algorithms and data structures.

The remainder of this chapter is organized as the following. Section 4.2 lays the mathematical foundations for presenting the algorithm. Section 4.2.1 introduces the assumed physical setting. Section 4.2.2 introduces the perturbation theory formalism that we use. Section 4.2.3 expands on the intuition about viewing matrix products as walking on a graph and introduces its connection to infinity norm, which will become useful in later developments. Section 4.2.4 introduces symmetric polynomials, which serve as the bedrock of our algorithms. Section 4.2.5 discusses cellular automaton from the perspective of existing literature and the differences and similarities between our construction and existing ones.

Section 4.3 further elaborates the content of Section 4.2 in the context of perturbation theory and derives an upper bound for the magnitude of r^{th} order term as a sum of walks in the space of reduced energy configurations. Section 4.3.1 builds on Section 4.2.1 to elaborate on the structure of $\mathbf{V}$. Section 4.3.2 builds on the perturbation theory outlined in Section 4.2.2 by applying the notions introduced in Section 4.3.1. Section 4.3.3 builds on the linear algebraic intuition described in Section 4.2.3 by incorporating it into the perturbation theory in Section 4.3.2. Section 4.3.4 carries the notion of walking among $\mathbf{H}$ eigenstates, which is introduced in Section 4.3.3, into the domain of energy configurations $\mathbf{c}$. Section 4.3.5 describes how to transform the sum over walks in energy configurations $\mathbf{c}$ to a sum over walks in *reduced* energy configurations $\tilde{\mathbf{c}}$ by using the symmetric polynomial defined in 4.2.4, see Lemma 4.3.6.

Section 4.4 is the main section introducing our algorithms for computing the upper bounds established in Section 4.3. Section 4.4.1 describes the algorithm used for constructing the cellular automaton given the physical setting. Section 4.4.2

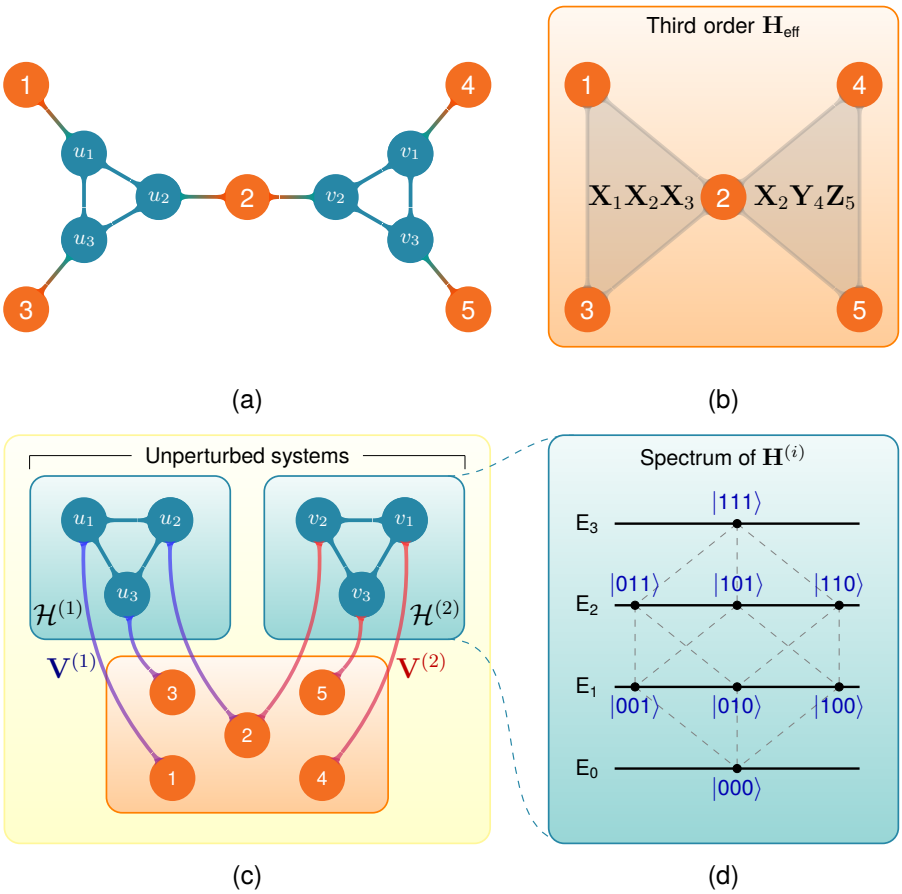
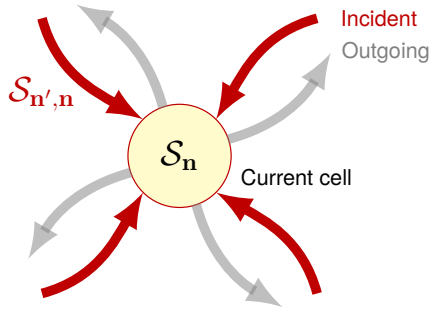
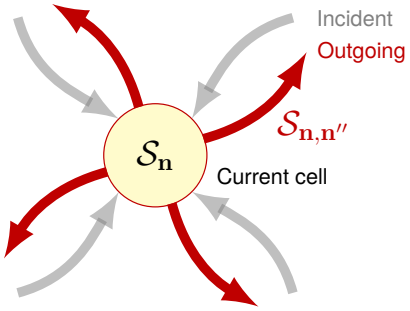


Fig. 4.3. A numerical example for demonstrating our algorithm estimating the perturbative error. (a): The 11-spin system constructed for testing. Each node corresponds to a spin-1/2 particle and each edge represents an interaction term in the Hamiltonian between two spins. (b): Effective Hamiltonian truncating at 3$^{\text{rd}}$ order perturbation theory. Here each triangle represents a 3-body interaction term. Using the perturbative expansion in Equation 4.1 we could show that the low-energy effective Hamiltonian truncated at 3$^{\text{rd}}$ order is $\mathbf{H}_{\text{eff}} = \alpha_1\mathbf{X}_1\mathbf{X}_2\mathbf{X}_3 + \alpha_2\mathbf{X}_2\mathbf{Y}_4\mathbf{Z}_5$ up to a constant energy shift. (c): Rearranging and partitioning the system in (a) according to the setting of perturbation theory used. Here each unperturbed system $\mathbf{H}^{(i)}$ consists of three ferromagnetically interacting spins (details in the long version). (d): Spectrum of each subsystem $\mathbf{H}^{(i)}$ in (a), $i \in \{1, 2\}$. Here each node represents an eigenstate of $\mathbf{H}^{(i)}$. Nodes on a same horizontal dashed line belong to the same energy subspace $\mathcal{P}_j$. There is an edge (u, v) iff $|\langle u|\mathbf{V}|v\rangle| \neq 0$. For example, if we consider this diagram as representing $\mathbf{H}^{(1)}$, since $\mathbf{V}^{(1)}|001\rangle_{u_1u_2u_3} \propto (|101\rangle + |011\rangle + |000\rangle)_{u_1u_2u_3}$ we connect the $|001\rangle$ with the nodes representing $|101\rangle$, $|011\rangle$ and $|000\rangle$.



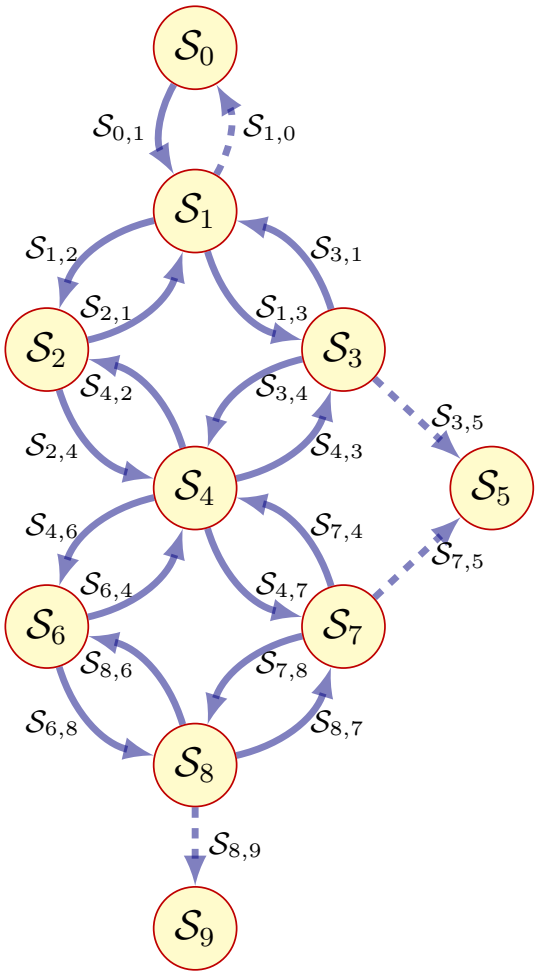
(a) Phase I



(b) Phase II

Cell	$\mathbf{n}$	$E(\mathbf{n})$	Subspace
$\mathcal{S}_0$	(2, 0, 0, 0)	E_0	$\mathcal{L}_-$
$\mathcal{S}_1$	(1, 1, 0, 0)	E_1	$\mathcal{L}_+$
$\mathcal{S}_2$	(0, 2, 0, 0)	$2E_1$	$\mathcal{L}_+$
$\mathcal{S}_3$	(1, 0, 1, 0)	E_2	$\mathcal{L}_+$
$\mathcal{S}_4$	(0, 1, 1, 0)	$E_1 + E_2$	$\mathcal{L}_+$
$\mathcal{S}_5$	(1, 0, 0, 1)	E_3	$\mathcal{L}_-$
$\mathcal{S}_6$	(0, 0, 2, 0)	$2E_2$	$\mathcal{L}_+$
$\mathcal{S}_7$	(0, 1, 0, 1)	$E_1 + E_3$	$\mathcal{L}_+$
$\mathcal{S}_8$	(0, 0, 1, 1)	$E_2 + E_3$	$\mathcal{L}_+$
$\mathcal{S}_9$	(0, 0, 0, 2)	$2E_3$	$\mathcal{L}_-$

(c)



(d)

Fig. 4.4. The cellular automaton generated for the example considered in Figure 4.3. Here each cell corresponds to an energy level of the unperturbed system $\mathbf{H} = \mathbf{H}^{(1)} + \mathbf{H}^{(2)}$. The sets of 4-tuples $\mathcal{S}_i$ and $\mathcal{S}_{i,j}$ at each cell and each directed edge store lists of 4-tuples $(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu)$. (a) and (b): Schematic diagrams for illustrating the two sequential steps executed when updating the state of each cell during an iteration. (c): A table listing the energy combinations $\mathbf{n}$, energy $E(\mathbf{n})$ and the subspace (low energy $\mathcal{L}_-$ or high energy $\mathcal{L}_+$) associated with each cell. (d): The cellular automaton constructed for the example considered in Figure 4.3 and Equation 4.71. Here the dashed lines corresponds to edges that go from a node in $\mathcal{L}_+$ to one in $\mathcal{L}_-$, which is only present in the automaton during the final step.

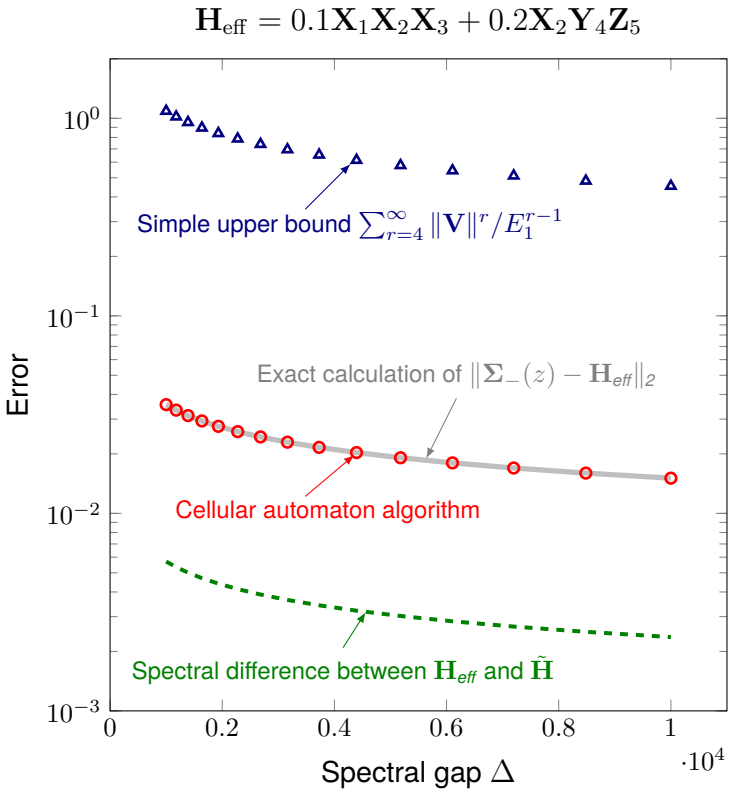


Fig. 4.5. Comparison between the upper bounds computed using the cellular automaton algorithm and the norm computed using (inefficient) explicit method. The "actual spectral error" in this plot shows the maximum difference between the eigenvalues of $\mathbf{H}_{\text{eff}}$ and their counterparts in $\widetilde{\mathbf{H}}$, which are the energies of its 2^N lowest eigenstates with $N = 5$ being the number of particles that $\mathbf{H}_{\text{eff}}$ acts on (Figure 4.3b). The actual spectral error is always lower than the error computed based on $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2$ because $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2 \leq \epsilon$ is only a *sufficient* condition that guarantees the spectral difference between $\widetilde{\mathbf{H}}$ and $\mathbf{H}_{\text{eff}}$ being within ϵ (see Theorem 1.3.1 and its variant Theorem 3.2.1).

describes the update rules for the cells. Section 4.4.3 shows the final algorithm for computing upper bounds of perturbative terms at arbitrary order r.

Section 4.5 shows a concrete example of a physical system and we conclude with Section 4.6, where we discuss the potential uses of our technique in a broader context of physical theories that require perturbative treatment. **Due to a large amount of symbols and notations introduced in this Chapter, we provide a glossary for these symbols in alphabetical order in Appendix C.**

4.2 Preliminaries

4.2.1 Basic setting

We consider the most general setting of perturbation theory, where we have an *unperturbed Hamiltonian* $\mathbf{H}$ with an energy gap Δ between its ground state subspace $\mathcal{L}_-$ and the rest of its spectrum which we denote as $\mathcal{L}_+$. Naturally in the eigenbasis of $\mathbf{H}$ we could write down $\mathbf{H}$ as a block diagonal operator:

$$\mathbf{H} = \begin{pmatrix} \mathbf{H}_+ & \\ & \mathbf{H}_- \end{pmatrix}. \tag{4.7}$$

Then we add a perturbation $\mathbf{V}$ to the unperturbed Hamiltonian. Here we assume $\|\mathbf{V}\|_2 < \Delta/2$. Here $\|\cdot\|_2$ is the 2-norm defined as $\|\mathbf{A}\|_2 = \max_{\||\psi\rangle\|=1} \|\mathbf{A}|\psi\rangle\|$. In the same basis we could write $\mathbf{V}$ as a block matrix

$$\mathbf{V} = \begin{pmatrix} \mathbf{V}_+ & \mathbf{V}_{+-} \\ \mathbf{V}_{-+} & \mathbf{V}_- \end{pmatrix}. \tag{4.8}$$

For a parameter z such that $|z| \ll \Delta$, define *operator valued resolvent* $\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$. Then like $\mathbf{H}$, $\mathbf{G}$ is also block diagonal in the eigenbasis of $\mathbf{H}$.

Suppose we are most concerned with the low-energy subspace of the perturbed Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$, which is spanned by all the eigenvectors of $\widetilde{\mathbf{H}}$ with eigen-

values that are less than $\Delta/2$. However, we do not require that the ground state of $\mathbf{H}$ be necessarily non-degenerate.

The unperturbed Hamiltonian $\mathbf{H}$ should correspond to some finite physical system with ℓ energy levels E_0, E_1, E_2, $\cdots$, $E_{\ell-1}$ with the corresponding eigenspaces which we denote as $\mathcal{P}_0$, $\mathcal{P}_1$, $\mathcal{P}_2$, $\cdots$, $\mathcal{P}_{\ell-1}$ and the respective projectors as $\mathbf{P}_0$, $\mathbf{P}_1$, $\mathbf{P}_2$, $\cdots$, $\mathbf{P}_{\ell-1}$.

Without loss of generality assume E_0, the ground state energy of $\mathbf{H}$, is zero. The energy values E_i do not have to be distinct or monotonically increasing but they should be separable into two subsets with one corresponding to the low-energy subspace $\mathcal{L}_- = \text{span}\{|E_j\rangle | E_j < \Delta/2\}$ and the other one corresponding to the rest of the spectrum $\mathcal{L}_+ = \text{span}\{|E_j\rangle | E_j > \Delta/2\}$.

Now let us consider a setting with m identical copies of such systems described by $\mathbf{H}$, each of which we call a *subsystem*. In this case all of the m subsystems are mutually non-interacting. The possible total energy values of the this m-copy system are thus simply linear combinations of energy levels of each subsystem. In essence, the spectrum of the m-copy system can be described by the set

$$\left\{ E = \sum_{i=0}^{\ell} n_i E_i \quad | \sum_{i=0}^{\ell} n_i = m, \quad n_i \in \mathbb{Z}, \quad 0 \leq n_i \leq m \right\}. \tag{4.9}$$

Let $\mathcal{H}$ be the Hilbert space where $\mathbf{H}$ dwells. As a notation we use $\mathcal{H}^{(i)}$, $i = 1, \cdots, m$, to denote the Hilbert space associated with the i^{th} subsystem. Let $\mathbf{H}^{(i)}$ be the Hamiltonian of the i^{th} subsystem. Correspondingly we introduce the notations for eigenvalues $E_j^{(i)}$, eigenspaces $\mathcal{P}_j^{(i)}$ spanned by eigenvectors $|\psi_{j,p}^{(i)}\rangle$ with p ranging from 1 to $\dim(\mathcal{P}_j^{(i)})$ and their projectors $P_j^{(i)}$ defined as

$$\mathbf{P}_j^{(i)} = \sum_{p=1}^{\dim\left(\mathcal{P}_j^{(i)}\right)} |\psi_{j,p}^{(i)}\rangle\langle\psi_{j,p}^{(i)}|. \tag{4.10}$$

where $|\psi_{j,p}^{(i)}\rangle$ represents the p^{th} degenerate eigenstate of $\mathbf{H}^{(i)}$ with energy E_j.

Now we further introduce perturbation $\mathbf{V}$ for each subsystem, by letting each of the subsystems interact with a common "bath" with Hilbert space $\mathcal{B}$, as illustrated in

Figure 1a of the main text. $\mathbf{V}$ contains a sum of terms $\mathbf{V}^{(i)}$ that couples the eigenspace $\mathcal{H}^{(i)}$ of the i^{th} unperturbed subsystem with the Hilbert space of the "bath" $\mathcal{B}$ by acting non-trivially on the joint space $\mathcal{H}^{(i)} \otimes \mathcal{B}$.

The "bath" by itself has its own internal dynamics governed by some Hamiltonian we write as $\mathbf{H}_{\mathcal{B}}$. This Hamiltonian describes interactions in $\mathcal{B}$ that are independent of each subspace $\mathcal{H}^{(i)}$. We point out that both the $\mathbf{H}^{(i)}$'s and $\mathbf{V}^{(i)}$'s act on the total Hilbert space $\tilde{\mathcal{H}} = \mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)} \otimes \cdots \otimes \mathcal{H}^{(m)} \otimes \mathcal{B}$ but only non-trivially on $\mathcal{H}^{(i)}$ for the $\mathbf{H}^{(i)}$'s and $\mathcal{H}^{(i)} \otimes \mathcal{B}$ for the $\mathbf{V}^{(i)}$'s. Like before we could also partition each of the local subspace $\mathcal{H}^{(i)}$ into low and high energy subspaces $\mathcal{L}_-^{(i)}$ and $\mathcal{L}_+^{(i)}$ such that $\mathcal{H}^{(i)} = \mathcal{L}_-^{(i)} \oplus \mathcal{L}_+^{(i)}$. Then the total Hilbert space can be written as $\tilde{\mathcal{H}} = (\mathcal{L}_- \oplus \mathcal{L}_+) \otimes \mathcal{B}$ where $\mathcal{L}_- = \mathcal{L}_-^{(1)} \otimes \mathcal{L}_-^{(2)} \otimes \cdots \otimes \mathcal{L}_-^{(m)}$ and $\mathcal{L}_+$ is the complement of $\mathcal{L}_-$ in the subspace $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)} \otimes \cdots \otimes \mathcal{H}^{(m)}$.

With definitions of subspaces in place, we define the unperturbed Hamiltonian $\mathbf{H}$ and the perturbation $\mathbf{V}$ as

$$\mathbf{H} = \sum_{i=1}^{m} \mathbf{H}^{(i)}, \qquad \mathbf{V} = \mathbf{H}_{\mathcal{B}} + \sum_{i=1}^{m} \mathbf{V}^{(i)}. \tag{4.11}$$

For each subsystem i, we assume that the perturbation $\mathbf{V}$ induces only transitions between $\mathcal{P}_j^{(i)}$ and $\mathcal{P}_k^{(i)}$ such that j and k differ by at most one. In other words, for any $i = 1, 2 \cdots, m$, we assume that the perturbation $\mathbf{V}$ be block tridiagonalizable in the eigenbasis of $\mathbf{H}$:

$$\mathbf{V}^{(i)} = \mathbf{1}_{\mathcal{H}^{(1)}} \otimes \mathbf{1}_{\mathcal{H}^{(2)}} \otimes \cdots \mathbf{1}_{\mathcal{H}^{(i-1)}} \otimes \begin{pmatrix} \mathbf{O}_{00}^{(i)} & \mathbf{O}_{01}^{(i)} & & & \\ \mathbf{O}_{10}^{(i)} & \mathbf{O}_{11}^{(i)} & \mathbf{O}_{12}^{(i)} & & \\ & \mathbf{O}_{21}^{(i)} & \ddots & \ddots & \\ & & \ddots & \mathbf{O}_{\ell-2,\ell-2}^{(i)} & \mathbf{O}_{\ell-2,\ell-1}^{(i)} \\ & & & \mathbf{O}_{\ell-1,\ell-2}^{(i)} & \mathbf{O}_{\ell-1,\ell-1}^{(i)} \end{pmatrix} \otimes \cdots$$

$$\cdots \otimes \mathbf{1}_{\mathcal{H}^{(i+1)}} \otimes \cdots \otimes \mathbf{1}_{\mathcal{H}^{(m)}}. \tag{4.12}$$

Here each block $\mathbf{O}_{jk}^{(i)}$ represents the transition driven by the perturbation $\mathbf{V}$ from states in the eigenspace $\mathcal{P}_j^{(i)}$ to those in $\mathcal{P}_k^{(i)}$. With further block permutation by grouping the blocks $\mathbf{O}_{jk}^{(i)}$ according to whether indices j and k correspond to $+$ or $-$ subspace, we could rewrite $\widetilde{\mathbf{H}}$ in the block form consistent with 4.7 and 4.8.

4.2.2 Perturbation theory

Let $\mathbf{\Pi}_-$ and $\mathbf{\Pi}_+$ be projectors onto the subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively. Then the block form of Equations 4.7 and 4.8 still holds for the definitions of $\mathbf{H}$ and $\mathbf{V}$ in Equation 4.11. More generally, any operator $\mathcal{O}$ can be written as the block form

$$\begin{pmatrix} \mathcal{O}_+ & \mathcal{O}_{+-} \\ \mathcal{O}_{-+} & \mathcal{O}_- \end{pmatrix}. \tag{4.13}$$

Our goal is to find a series expansion that approximates the low-energy effective Hamiltonian of the perturbed system $\widetilde{\mathbf{H}}_-$. In Section 4.2.1 we defined the operator-valued resolvent $\mathbf{G}(z) = (z\mathbf{I} - \mathbf{H})^{-1}$. We could similarly define operator-valued resolvent $\tilde{\mathbf{G}}(z) = (z\mathbf{I} - \widetilde{\mathbf{H}})^{-1}$ for $z \ll \Delta$ where $\mathbf{I}$ is the identity acting on $\tilde{\mathcal{H}}$. We could relate $\tilde{\mathbf{G}}$ with $\mathbf{G}$ by $\tilde{\mathbf{G}} = (\mathbf{G}^{-1} - \mathbf{V})^{-1}$, which gives rise to a Taylor expansion

$$\tilde{\mathbf{G}} = \mathbf{G}(\mathbf{I} - \mathbf{V}\mathbf{G})^{-1} = \mathbf{G} + \mathbf{G}\mathbf{V}\mathbf{G} + \mathbf{G}\mathbf{V}\mathbf{G}\mathbf{V}\mathbf{G} + \cdots \tag{4.14}$$

Recall from Section 1.3 the central object of our concern, namely the *self-energy expansion* $\mathbf{\Sigma}_-(z) = zI - (\widetilde{\mathbf{G}}_-(z))^{-1}$. Applying 4.13 and 4.14 on $\mathbf{\Sigma}_-(z)$ leads to

$$\begin{aligned} \mathbf{\Sigma}_-(z) &= \mathbf{H}_- + \mathbf{V}_- + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} + \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\mathbf{G}_+\mathbf{V}_{+-} + \cdots \\ &= \mathbf{H}_- + \mathbf{V}_- + \sum_{r=2}^{\infty} \mathbf{T}_r. \end{aligned} \tag{4.15}$$

The self-energy $\mathbf{\Sigma}_-(z)$ is important for approximating the low-energy effective Hamiltonian $\widetilde{\mathbf{H}}_-$. The following theorem makes this intuition precise.

Theorem 4.2.1 (Theorem 1.3.1 restated, also adapted from [7], [8]) *Given a Hamiltonian* $\widetilde{\mathbf{H}} = \mathbf{H}+\mathbf{V}$. *Suppose* $\|\mathbf{V}\|_2 \leq \Delta/2$ *with* Δ *being the spectral gap between the ground and the first excited state of* $\mathbf{H}$. *If there exists a Hamiltonian* $\mathbf{H}_{\mathit{eff}}$ *whose energies are contained in the interval* $[a, b]$ *and some real constant* $\epsilon > 0$ *such that* $a < b < \Delta/2 - \epsilon$ *and for any* $z \in [a - \epsilon, b + \epsilon]$,

$$\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\mathit{eff}}\|_2 \leq \epsilon, \tag{4.16}$$

then the j^{th} *eigenvalue* $\tilde{\lambda}_j$ *of* $\widetilde{\mathbf{H}}_-$ *and the corresponding* j^{th} *eigenvalue of* $\mathbf{H}_{\mathit{eff}}$ *differ by at most* ϵ, *for any appropriate range of* j *values.*

Most uses of perturbation theory involve truncating the perturbative expansion 4.15 to a specific order to obtain an effective Hamiltonian $\mathbf{H}_{\text{eff}}$ that approximates the exact solution. Theorem 4.2.1 is valuable in the sense that it establishes a connection between the magnitude of the error term $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2$ and the quality of $\mathbf{H}_{\text{eff}}$ as an approximation to $\widetilde{\mathbf{H}}$, modulo certain conditions that are clearly satisfied by our assumed physical setting described in Section 4.2.1. The task of evaluating the quality of perturbative approximation is then reduced to the task of estimating the perturbative error $\|\mathbf{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|_2$. More specifically, our goal is to find a tight yet efficiently computable upper bound for the norm of the r^{th} order term $\mathbf{T}_r$ which is

$$\mathbf{T}_r \equiv \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{r-2}\mathbf{G}_+\mathbf{V}_{+-}, \tag{4.17}$$

Obviously one can obtain a crude bound by triangle inequality and submultiplicativity of operator norm (namely $\|\mathbf{AB}\|_2 \leq \|\mathbf{A}\|_2 \cdot \|\mathbf{B}\|_2$)

$$\|\mathbf{T}_r\|_2 \leq \|\mathbf{V}\|_2^r \cdot \|\mathbf{G}_+\|_2^{r-1}. \tag{4.18}$$

However, as we will demonstrate with a concrete example in Section 4.5, this does not serve as a bound tight enough to capture the true magnitude of $\|\mathbf{T}_r\|_2$. In order to find a tighter bound for $\|\mathbf{T}_r\|_2$, an extreme would be to explicitly form $\mathbf{T}_r$ and compute $\|\mathbf{T}_r\|_2$ directly. But the computation cost is evidently exponential in the size of the system. For the remainder of the Chapter we present a middle-ground

possibility where a tighter bound than $\|\mathbf{V}\|_2^r \cdot \|\mathbf{G}_+\|_2^{r-1}$ can be obtained by efficient computation. We show that in certain cases the bound obtained is even equal to the value of $\|\mathbf{T}_r\|_2$, providing evidence that significant improvement over our approach for general settings is likely difficult.

4.2.3 Matrix product, walks on graphs and the infinity norm

In this section we note a few intuitions concerning matrices that will be instrumental to our later discussions. We start by pointing out the connection between matrix products and walks on graphs. An $N \times N$ matrix $\mathbf{A} = \sum_{i,j} a_{ij}|i\rangle\langle j|$ could be considered as a weighted directed graph on N nodes with the edge from i to j having weight a_{ij}. In other words, each element a_{ij} signifies the "weight" of a walk $i \to j$. If we consider the product between $\mathbf{A}$ and another $N \times N$ matrix $\mathbf{B} = \sum_{i,j} b_{ij}|i\rangle\langle j|$, the (i,j) element of the product $\mathbf{AB}$ is $(\mathbf{AB})_{ij} = \sum_k a_{ik}|i\rangle\langle k| \cdot b_{kj}|k\rangle\langle j|$, which is a 2-step walk $i \to j \to k$. One could think of our central object $\mathbf{T}_r$ defined in Equation 4.17 as a collection of r-step walks in the space of $\mathbf{H}$ eigenstates. We will make this notion precise later.

Much of our arguments in our proofs of correctness for the algorithms will be based on ∞-norm, instead of 2-norm, of matrices. As a simple reminder, the ∞-norm of an $m \times n$ matrix $\mathbf{A}$ is defined as $\|\mathbf{A}\|_\infty = \max\limits_{1 \le i \le m} \sum_{j=1}^{n} |a_{ij}|$, which is simply the maximum absolute row sum of the matrix. We will be using the following properties of the infinity norm of matrices:

1. For any matrices $\mathbf{A}$ and $\mathbf{B}$ of compatible dimensions, $\|\mathbf{A} + \mathbf{B}\|_\infty \le \|\mathbf{A}\|_\infty + \|\mathbf{B}\|_\infty$;

2. For any matrices $\mathbf{A}$ and $\mathbf{B}$ of compatible dimensions, $\|\mathbf{AB}\|_\infty \le \|\mathbf{A}\|_\infty \cdot \|\mathbf{B}\|_\infty$;

3. $\|\mathbf{A} \otimes \mathbf{1}\|_\infty = \|\mathbf{A}\|_\infty$ where $\mathbf{1}$ is an identity matrix of any finite dimension. Similarly $\|\mathbf{1} \otimes \mathbf{A}\|_\infty = \|\mathbf{A}\|_\infty$;

4. If $\mathbf{A}$ is a block matrix and let $\mathbf{A}_{ij}$ be the (i,j) block[1], then

$$\|\mathbf{A}\|_\infty \leq \max_i \sum_j \|\mathbf{A}_{ij}\|_\infty;$$

5. For a Hermitian matrix $\mathbf{A}$, we have $\|\mathbf{A}\|_2 \leq \|\mathbf{A}\|_\infty$. This follows from $\|\mathbf{A}\|_2^2 \leq \|\mathbf{A}\|_1 \cdot \|\mathbf{A}\|_\infty$ and $\|\mathbf{A}\|_1 = \|\mathbf{A}\|_\infty$ for Hermitian matrices.

Here Property 5 is useful because it ties directly to 2-norm, which has a natural connection to the spectrum of the matrix and is more commonly used for characterizing the magnitude of perturbative error $\mathbf{T}_r$ at any order r. Our algorithms, on the other hand are intended for computing upper bounds to $\|\mathbf{T}_r\|_\infty$. Property 5 thus guarantees that the upper bounds computed for $\|\mathbf{T}_r\|_\infty$ also serve as upper bounds to $\|\mathbf{T}_r\|_2$.

We prefer to use infinity norm in the context of this work because of its natural connection to the element-wise or block-wise structure of a matrix. Drawing on the connection mentioned in the opening paragraph, consider the powers of a block matrix $\mathbf{A}$, namely A^n. Following the notation in Property 4, let $\mathbf{A}_{ij}$ be the (i,j) block. Assume $\mathbf{A}$ is an $k \times k$ block matrix. If we think of the matrix $\mathbf{A}$ as a directed weighted graph on k nodes where each edge going from node i to j is associated with "weight" $\mathbf{A}_{ij}$, then the (i,j) block of $\mathbf{A}^n$ essentially is a sum over contributions from all n-step walks $i_0 \to i_1 \to i_2 \to \cdots \to i_n$ on the graph of $\mathbf{A}$ that starts from $i_0 = i$ and ends at $i_n = j$. Each one of such n-step walk contributes a term $\mathbf{A}_{i_0 i_1}\mathbf{A}_{i_1 i_2}\cdots\mathbf{A}_{i_{n-1} i_n}$ to the (i,j) block of $\mathbf{A}^n$. Hence if use $(\mathbf{A}^n)_{ij}$ to denote the (i,j) block of $\mathbf{A}^n$,

$$(\mathbf{A}^n)_{ij} = \sum_{i_1, i_2, \cdots, i_{n-1}} \mathbf{A}_{i i_1}\mathbf{A}_{i_1 i_2}\cdots\mathbf{A}_{i_{n-2} i_{n-1}}\mathbf{A}_{i_{n-1} j}. \tag{4.19}$$

Using Property 1, 2, 4 and 5 of infinity norm on Equation 4.19 we could find an upper bound

$$\|\mathbf{A}^n\|_2 \leq \max_{i=1,\cdots,k} \sum_{j=1}^{k} \sum_{i_1, i_2, \cdots, i_{n-1}} \|\mathbf{A}_{i i_1}\|_\infty \cdot \|\mathbf{A}_{i_1 i_2}\|_\infty \cdots \|\mathbf{A}_{i_{n-2} i_{n-1}}\|_\infty \cdot \|\mathbf{A}_{i_{n-1} j}\|_\infty. \tag{4.20}$$

[1]In our notation (i,j) block means the block on the i^{th} row and j^{th} column.

Equation 4.20 underlies the basic intuition of our approach in finding a tight upper bound to $\|\mathbf{T}_r\|_2$. Similar to Equation 4.20, $\mathbf{T}_r = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{r-2}\mathbf{G}_+\mathbf{V}_{+-}$ also contains a basic structure of powering the matrix $\mathbf{G}_+\mathbf{V}_+$. As later discussion would reveal, in the context of bounding $\|\mathbf{T}_r\|_\infty$ the walks over which the right hand side of Equation 4.20 sums over correspond to sequences of transitions among eigenstates of the unperturbed Hamiltonian $\mathbf{H}$. However, note that the sum over $i_1, i_2, \cdots, i_{n-1}$ in Equation 4.20 contains an exponential number of terms in n due to the permutation of indices, which means any naive algorithm that computes the right hand side of Equation 4.20 will likely be inefficient. We introduce a mathematical tool in the next section to help with this inefficiency due to combinatorics.

4.2.4 Symmetric polynomials

Symmetric polynomials are used in our algorithms as a fundamental data structure to address the combinatorics of arbitrary-order virtual transitions in the perturbative expansion. We start with a few definitions. Any *monomial* in n variables x_1, x_2, $\cdots$, x_n can be written as $x_1^{a_1}\cdots x_n^{a_n}$ where the exponents $\alpha_i \in \{0, 1, 2, \cdots\}$. Writing $\mathbf{a} = (a_1, a_2, \cdots, a_n)$ and $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ gives the abbreviated notation $\mathbf{x}^\mathbf{a} = x_1^{a_1}\cdots x_n^{a_n}$.

Definition 4.2.1 (Monomial symmetric polynomial) *The* monomial symmetric polynomial $m_\mathbf{a}(\mathbf{x})$ *is defined as the sum of all monomials* $\mathbf{x}^\mathbf{a}$ *where* $\mathbf{a}$ *ranges over all distinct permutations of elements in* $\mathbf{a} = (a_1, a_2, \cdots, a_n)$. *Here* $\mathbf{a}$ *can be thought of as a partition of an integer* $K = \sum_{i=1}^n a_i$ *and we say* $\mathbf{a}$ *is the* partition *of* $m_\mathbf{a}(\mathbf{x})$.

Note that by definition, a monomial symmetric polynomial is invariant with respect to the ordering of elements in the partition. For convenience we impose the following restrictions to the representations of partitions, which we call *reduced partition*. From here on we will only use the reduced partition to uniquely describe a monomial symmetric polynomial.

Definition 4.2.2 (Reduced partition) *For an n-variable monomial symmetric polynomial $m_{\mathbf{a}}(\mathbf{x})$, let k be the number of nonzero elements in $\mathbf{a}$. Then we define the* reduced partition $\mathbf{b}$ *of $m_{\mathbf{a}}(\mathbf{x})$ to be a k-dimensional vector formed by taking all the k nonzero elements of $\mathbf{a}$ and order them in non-descending order* i.e. $b_1 \leq b_2 \leq \cdots \leq b_k$.

There is a certain combinatorial intuition associated with monomial symmetric polynomials which is important in the context of later discussions. For instance consider $m_{(1,2,3)}(a,b,c) = ab^2c^3+ba^2c^3+ac^2b^3+ca^2b^3+bc^2a^3+cb^2a^3$. As an analogy, we could think of each variable a, b, c as a bucket of coins and each term in $m_{(1,2,3)}(a,b,c)$ as a result of flipping the coins in the three buckets one at a time such that in the end one bucket gets 1 coin flips, one gets 2 coin flips and the other gets 3. Each coin flip does not have to be on different coins. For example the first term, ab^2c^3, corresponds to the case where we administer 1 coin flip in bucket A, 2 coin flips in bucket B and 3 in C.

Another feature of monomial symmetric polynomial that we use is its compactness in representation. For $\mathbf{b}$ such that $\sum_{i=1}^{|\mathbf{b}|} = r$, $m_{\mathbf{b}}(x_1, \cdots, x_n)$ contains $O(n^r)$ terms, while all the information for generating these terms can be condensed to $\mathbf{b}$, a k-element vector. As is shown in Appendix D, for a fixed partition $\mathbf{b}$, evaluating $m_{\mathbf{b}}(x_1, x_2, \cdots, x_n)$ takes $O(r!n)$ time. In our context r is the order of perturbation, which is assumed to be fixed. Hence the cost of evaluating symmetric polynomials scales *linearly* as the number of variables (or in our context the number of unperturbed subsystems).

4.2.5 Cellular automata

A *cellular automaton* (CA) is typically defined as a collection of finite-state machines called *cells* that are positioned on a grid of any finite dimension. Each cell in the grid also has a defined set of other cells as its *neighborhood*. The initial configuration of the automaton is specified by assigning states to each cell in the grid. The cells evolve together in discrete time steps, each time updating the state of each cell

by a rule that is identical for each cell and does not change over time. During each time step, the rule determines the new state of each cell in terms of the current state of the cell and the states of the cells in its neighborhood.

While the initially proposed CA constructions adhere strictly to the definitions above, CA constructions that deviate from the above definitions abound. This has significantly added flexibility in the use of the terminology. For example,

- The states of cells need not be discrete; continuous-valued CAs in two-dimensions have been explored [174];
- The grid that joins the cells could be more than two-dimensional [175];
- More generally, the states of the cells do not necessarily have to be single numbers, but could also be data structures [174].

In this work we construct CAs that admit all three variations, namely CAs with cells connected in form of a (possibly high dimensional) grid and cell states that consist of data structures designed to specifically suit our purpose. However, our construction retains some typical features of cellular automata:

- The update rules are *local* in the sense that the states of the cells are only dependent on their neighbors;
- The update rules are *homogeneous* in that they are identical and time independent for all cells;
- The states of the cells are updated *in parallel* to produce a new generation.

An important problem concerning the theory of cellular automata is "What higher-level descriptions of information processing in cellular automata can be given?" [176]. There have been prior works [177] on CA constructions that are strongly based on analogues with conventional serial-processing computers. However, information processing in cellular automata occurs in a fundamentally distributed and parallel fashion. In this sense, the CAs constructed in this work perform computations in ways

that departs from conventional serial computer models: to obtain an upper bound to the norm of m^{th} order perturbative term, we evolve the CA for m evolutions and glean results from the states of a specific subset of cells.

4.3 Upper bounds for arbitrary order perturbation theory

4.3.1 Structure of the perturbation

In our basic setting we have assumed the perturbation $\mathbf{V}$ be block tridiagonalizable with respect to subspaces of $\mathcal{H}^{(i)}$, see Equation 4.12. Each block $\mathbf{O}_{jk}^{(i)}$ by itself has a block structure. Each $\mathbf{O}_{jk}^{(i)}$ is a $\dim(\mathcal{P}_j^{(i)}) \times \dim(\mathcal{P}_k^{(i)})$ array of operators $\mathbf{B}_{pq,jk}^{(i)}$ that only acts on $\mathcal{B}$. Explicitly,

$$\mathbf{O}_{jk}^{(i)} = \begin{pmatrix} \mathbf{B}_{11,jk}^{(i)} & \mathbf{B}_{12,jk}^{(i)} & \cdots & \mathbf{B}_{1K,jk}^{(i)} \\ \mathbf{B}_{21,jk}^{(i)} & \mathbf{B}_{22,jk}^{(i)} & \cdots & \mathbf{B}_{2K,jk}^{(i)} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{J1,jk}^{(i)} & \mathbf{B}_{J2,jk}^{(i)} & \cdots & \mathbf{B}_{JK,jk}^{(i)} \end{pmatrix} \tag{4.21}$$

where for convenience we define $J = \dim(\mathcal{P}_j^{(i)})$ and $K = \dim(\mathcal{P}_k^{(i)})$. Here $\mathbf{B}_{pq,jk}^{(i)}$ describes the action on $\mathcal{B}$ that is coupled with transition from the p^{th} degenerate state in $\mathcal{P}_j^{(i)}$ to the q^{th} degenerate state in $\mathcal{P}_k^{(i)}$.

The following definitions of quantities will become instrumental to our further development in this work.

Definition 4.3.1 (Scalar quantity ω) *Let ω be an upper bound to the norm of the components in $\mathbf{V}$ such that*

$$\omega \geq \|\mathbf{H}_{\mathcal{B}}\|_\infty + \max_{\substack{j=0,\cdots,\ell \\ i=1,\cdots,m}} \|\mathbf{O}_{jj}^{(i)}\|_\infty. \tag{4.22}$$

Definition 4.3.2 (Vector $\boldsymbol{\lambda}$) *Let λ_i be an upper bound to the norms of the matrix elements in the off-diagonal blocks $\mathbf{O}_{jk}^{(i)}$ (i.e. the blocks with j and k differing by one). In other words,*

$$\lambda_i = \max_{\substack{j,k=0,\cdots,\ell,\\ j\neq k}} \max_{\substack{p=1,\cdots,J,\\ q=1,\cdots,K}} \|\mathbf{B}_{pq,jk}^{(i)}\|_\infty. \tag{4.23}$$

For convenience we define the vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \cdots, \lambda_m)$.

Definition 4.3.3 (Matrix M) *For each block $\mathbf{O}_{jk}^{(i)}$ as defined in 4.12 and 4.21, let $M_{jk}^{(i)}$ be the maximum number of nonzero blocks per row in $\mathbf{O}_{jk}^{(i)}$. In precise terms,*

$$M_{jk}^{(i)} = \max_{p=1,\cdots,J} Card\{\mathbf{B}_{pq,jk}^{(i)}, q = 1, \cdots, K | \|\mathbf{B}_{pq,jk}^{(i)}\|_\infty \neq 0\} \tag{4.24}$$

where $Card\{\cdot\}$ is the size of a set. Furthermore, let $\mathbf{M}$ be an $\ell \times \ell$ matrix such that $M_{jk} = \max_{i=1,\cdots,m} M_{jk}^{(i)}$.

Informally, λ_i characterizes the "strength" of perturbation $\mathbf{V}$ acting on the subsystem $\mathcal{H}_i$ and causing a transition, while $M_{jk}^{(i)}$ characterizes the combinatorial aspect of $\mathbf{V}^{(i)}$ inducing transitions between eigenstates in the subspaces $\mathcal{P}_j$ and $\mathcal{P}_k$. Furthermore, M_{jk} represents the maximum possible ways, among all subsystems i, in which an unperturbed eigenstate in a subspace $\mathcal{P}_j^{(i)}$ can be transformed into an eigenstate in $\mathcal{P}_k^{(i)}$ via the action of $\mathbf{V}^{(i)}$. From a more linear algebraic perspective, it is the maximum row sparsity of the $\mathbf{O}_{jk}^{(i)}$ blocks among all subsystems.

4.3.2 Structure of terms at any order

The quantity $\mathbf{T}_r = \mathbf{V}_{-+}\mathbf{G}_+(\mathbf{V}_+\mathbf{G}_+)^{r-2}\mathbf{V}_{+-}$ is a string of matrices multiplied sequentially and we will consider finding upper bounds for the norm of each successively longer substring that starts with the first matrix $\mathbf{V}_{-+}$. By definition of block structures introduced in Equations 4.12 and 4.21, in the general setting described in Figure 1 of the main text we could express $\mathbf{V}_{-+}$ in terms of the finest block division $\mathbf{B}_{pq,jk}^{(i)}$ as

$$\mathbf{V}_{-+} = \sum_{i=1}^{m} \sum_{j:\mathcal{P}_j^{(i)}\subseteq\mathcal{L}_-^{(i)}} \sum_{k:\mathcal{P}_k^{(i)}\subseteq\mathcal{L}_+^{(i)}} \sum_{p:|\psi_{j,p}^{(i)}\rangle\in\mathcal{P}_j^{(i)}} \sum_{q:|\psi_{k,q}^{(i)}\rangle\in\mathcal{P}_k^{(i)}} \mathbf{B}_{pq,jk}^{(i)} \otimes |\psi_{j,p}^{(i)}\rangle\langle\psi_{k,q}^{(i)}| \tag{4.25}$$

where we recall that the operators $\mathbf{B}^{(i)}_{pq,jk}$ are defined in Equation 4.21 and the states $|\psi^{(i)}_{j,p}\rangle$ are defined in 4.10.

Following Equation 4.25 we could also express $\mathbf{G}_+$, $\mathbf{V}_+$ and $\mathbf{V}_{+-}$ in terms of blocks $\mathbf{B}^{(i)}_{pq,jk}$ and unperturbed eigenstates $|\psi^{(i)}_{j,p}\rangle$. Starting from $\mathbf{G}_+(z) = \mathbf{\Pi}_+(z\mathbf{I} - \mathbf{H})^{-1}\mathbf{\Pi}_+$, before expanding $\mathbf{G}_+$ we introduce the following notions of *energy combination* and *energy configuration* of a given eigenstate of $\mathbf{H}$. These notions are also important in our further algorithmic development.

Definition 4.3.4 (Energy configuration) *For an eigenstate $|\psi\rangle$ of $\mathbf{H}$ where the energy of each subsystem $\mathbf{H}^{(i)}$ is $E^{(i)} = \langle\psi|\mathbf{H}^{(i)}|\psi\rangle \in \{E_0, E_1, \cdots, E_\ell\}$. We define the* energy configuration *of the eigenstate $|\psi\rangle$ as a vector $\mathbf{c} \in \{0, 1, \cdots, \ell\}^m$ with each element $\mathbf{c}_i$ be such that $E^{(i)} = E_{c_i}$. We use the notation $\mathbf{c}(|\psi\rangle)$ to refer to the energy configuration of $|\psi\rangle$.*

Definition 4.3.5 (Energy combination) *Given an energy configuration $\mathbf{c}$, for each energy level j ranging from 0 to $\ell - 1$, let n_j be the number of subsystems with energy j. In other words $n_j = Card\{i = 1, \cdots, m | c_i = j\}$ where $Card\{\cdot\}$ is the cardinality of a set. Then we define the* energy combination *of the energy configuration $\mathbf{c}$ as $\mathbf{n}(\mathbf{c}) = (n_1, n_2, \cdots, n_\ell) \in \{1, 2, \cdots, m\}^\ell$. Conversely, let $\mathcal{C}(\mathbf{n}) = \{\mathbf{c} | \forall j \in \{0, \cdots, \ell\}, \sum_{i:c_i=j} 1 = n_j\}$ be the set of energy configuration that gives rise to a given energy combination $\mathbf{n}$.*

Informally one could think of $\mathbf{n}$ as representing the eigenstates of $\mathbf{H}$ in a "number basis". Then $G_+(z)$ can be expressed as

$$\mathbf{G}_+(z) = \sum_{\mathbf{n}\in\mathcal{N}_+} \frac{1}{z - E(\mathbf{n})} \sum_{\mathbf{c}\in\mathcal{C}(\mathbf{n})} \mathbf{P}(\mathbf{c}) \tag{4.26}$$

where $E(\mathbf{n}) = \sum_{j=1}^{\ell} n_j E_j$ is the total energy of the current energy combination. $\mathcal{N}_+ = \{\mathbf{n} | E(\mathbf{n}) > \Delta/2\}$ is the set of energy combination that correspond to an eigenstate of $\mathbf{H}$ in $\mathcal{L}_+$. Similarly we could also define $\mathcal{N}_- = \{\mathbf{n} | E(\mathbf{n}) < \Delta/2\}$. $\mathbf{P}(\mathbf{c}) = \bigotimes_{j=1}^{m} \mathbf{P}^{(j)}_{c_j}$ is the projector onto the subspace of each subsystem as described

by the energy configuration $\mathbf{c}$. Each of the projector $\mathbf{P}_{c_j}^{(j)}$ could be further expressed as projectors onto individual eigenstates by (4.10).

The expression for V_+ in terms of blocks $\mathbf{B}_{pq,jk}^{(i)}$ and unperturbed eigenstates $|\psi_{j,p}^{(i)}\rangle$ can be obtained by replacing $\mathcal{L}_-^{(i)}$ in the summation over j in (4.25) by $\mathcal{L}_+^{(i)}$. Similarly, the expression for V_{+-} can be obtained by replacing $\mathcal{L}_-^{(i)}$ in the summation over j in (4.25) by $\mathcal{L}_+^{(i)}$ and at the same time replacing $\mathcal{L}_+^{(i)}$ in the summation over k in (4.25) by $\mathcal{L}_-^{(i)}$.

4.3.3 Walk in the space of unperturbed eigenstates

With the notation $P(\mathbf{c})$ introduced in Equation 4.26 we could express $\mathbf{\Pi}_-$ and $\mathbf{\Pi}_+$ explicitly as

$$\mathbf{\Pi}_- = \sum_{\mathbf{n}\in\mathcal{N}_-} \sum_{\mathbf{c}\in\mathcal{C}(\mathbf{n})} \mathbf{P}(\mathbf{c}), \qquad \mathbf{\Pi}_+ = \sum_{\mathbf{n}\in\mathcal{N}_+} \sum_{\mathbf{c}\in\mathcal{C}(\mathbf{n})} \mathbf{P}(\mathbf{c}). \tag{4.27}$$

Combining Equation (4.27) with the definitions of $\mathbf{P}_j^{(i)}$ in Equation (4.10) we could see that the term

$$\mathbf{T}_r = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{r-2}\mathbf{G}_+\mathbf{V}_{+-}$$

for any $r \geq 3$ consists of products of $\mathbf{B}_{pq,jk}^{(i)} \otimes |\psi_{j,p}^{(i)}\rangle\langle\psi_{k,q}^{(i)}|$ with each term $|\psi_{j,p}^{(i)}\rangle\langle\psi_{k,q}^{(i)}|$ multiplied together forming a sequence of virtual transitions

$$\underbrace{|\phi^{(0)}\rangle\langle\phi^{(1)}|}_{\mathbf{V}_{-+}} \cdot \underbrace{|\phi^{(1)}\rangle\langle\phi^{(1)}|}_{\mathbf{G}_+} \cdot \underbrace{|\phi^{(1)}\rangle\langle\phi^{(2)}|}_{\mathbf{V}_+} \cdot \underbrace{|\phi^{(2)}\rangle\langle\phi^{(2)}|}_{\mathbf{G}_+} \cdots \underbrace{|\phi^{(r-1)}\rangle\langle\phi^{(r)}|}_{\mathbf{V}_{+-}} \tag{4.28}$$

that corresponds to a walk among the eigenstates of $\mathbf{H}$. For convenience in the subsequent discussions we temporarily condense all the subscripts j, p and superscript (i) of the state $|\psi_{j,p}^{(i)}\rangle$ into a single-number superscript. To avoid confusion with the superscript notation in $|\psi_{j,p}^{(i)}\rangle$ we use ϕ instead of ψ. The superscript for ϕ indicates the step of a walk while the superscript for ψ indicates the subsystem. We will only use $|\phi^{(i)}\rangle$ notation when referring to a generic walk among eigenstates of $\mathbf{H}$. Here in Equation (4.28) the operators indicated under the brackets "$\underbrace{\quad}$" are the operators

that contributes the respective projector $|\cdot\rangle\langle\cdot|$ in $\mathbf{T}_r$. We formally define such walk in the context of bounding $\|\mathbf{T}_r\|_2$ as the following.

Definition 4.3.6 (Walk in the space of H eigenstates) *We define an* r-step walk in the space of $\mathbf{H}$ eigenstates *as a sequence of unperturbed eigenstates* $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle \to \cdots \to |\phi^{(r)}\rangle$ *such that*

$$\begin{aligned} &|\phi^{(0)}\rangle \in \mathcal{L}_-, \quad |\phi^{(r)}\rangle \in \mathcal{L}_- \\ &|\phi^{(i)}\rangle \in \mathcal{L}_+, \quad i = 1, \cdots, r-1. \end{aligned} \tag{4.29}$$

In addition, we require that $\|\langle\phi^{(i)}|\mathbf{V}|\phi^{(i+1)}\rangle\| \neq 0$ *for any* $i = 0, 1, \cdots, r-1$. *Let* $E^{(i)}$ *be the energy of* $|\phi^{(i)}\rangle$, *namely* $E^{(i)} = \langle\phi^{(i)}|\mathbf{H}|\phi^{(i)}\rangle$.

Definition 4.3.6 is laid out specifically for enumerating terms in $\mathbf{T}_r$. The following lemma describes the explicit connection between the r^{th} order perturbative term $\mathbf{T}_r$ and the r-step walk in Definition 4.3.6.

Lemma 4.3.1 *For an* r*-step walk* $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle \to \cdots \to |\phi^{(r)}\rangle$, *let* $\mathbf{B}^{(i)}$ *be the* $\mathbf{B}^{(i)}_{pq,jk}$ *block in* $\mathbf{V}$ *(Equation 4.12 and 4.21) associated with the transition* $|\phi^{(i-1)}\rangle \to |\phi^{(i)}\rangle$. *In other words*[2], $\mathbf{B}^{(i)} \otimes |\phi^{(i-1)}\rangle\langle\phi^{(i)}| = \mathbf{B}^{(i)}_{pq,jk} \otimes |\psi^{(i)}_{j,p}\rangle\langle\psi^{(i)}_{k,q}|$. *Then*

$$\begin{aligned} \mathbf{T}_r = \sum_{|\phi^{(0)}\rangle \in \mathcal{L}_-} \sum_{|\phi^{(r)}\rangle \in \mathcal{L}_-} {\sum}' \mathbf{B}^{(1)} \cdot \frac{1}{|z - E^{(1)}|} \cdot \mathbf{B}^{(2)} \cdot \frac{1}{|z - E^{(2)}|} \cdots \\ \cdots \frac{1}{|z - E^{(r-1)}|} \cdot \mathbf{B}^{(r)} \otimes |\phi^{(0)}\rangle\langle\phi^{(r)}| \end{aligned} \tag{4.30}$$

where Σ' *sums over all* r*-step walks in the space of* $\mathbf{H}$ *eigenstates, as in Definition 4.3.6, but restricted to a* fixed *pair of* $|\phi^{(0)}\rangle$ *and* $|\phi^{(r)}\rangle$.

Proof In Section 4.2.3 we interpret powers of block matrices as walks on a weighted directed graph with each edge carrying a "weight" that is a block. Applying this intuition to the block partitioning of the perturbation $\mathbf{V}$ introduced in Section 4.3.1,

[2]To avoid confusion with the (i) superscripts we use $\mathbf{B}$ instead of $\mathbf{B}$. Here the superscript (i) of $\mathbf{B}^{(i)}$ stands for the i^{th} step in the walk, while superscript (i) of $\mathbf{B}^{(i)}_{pq,jk}$ represents the i^{th} subsystem.

we could see that $\mathbf{T}_r$ is also a block matrix of $\dim(\mathcal{L}_-) \otimes \dim(\mathcal{L}_-)$ blocks with the (i,j) block being the sum over all of the contributions from walks in the space of $\mathbf{H}$ eigenstates (Definition 4.3.6) that start from the i^{th} low energy eigenstate and end at the j^{th} low energy eigenstate. With $|\phi^{(0)}\rangle$ being the i^{th} low energy eigenstate and $|\phi^{(r)}\rangle$ being the j^{th}, one could see that a term in $\mathbf{T}_r$ corresponding to a walk $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle \to \cdots \to |\phi^{(r)}\rangle$ takes the form

$$\underbrace{(\mathbf{B}^{(1)} \otimes |\phi^{(0)}\rangle\langle\phi^{(1)}|)}_{\mathbf{V}_{-+}} \cdot \underbrace{\left(\frac{1}{z - E^{(1)}}|\phi^{(1)}\rangle\langle\phi^{(1)}|\right)}_{\mathbf{G}_+} \cdot \underbrace{(\mathbf{B}^{(2)} \otimes |\phi^{(1)}\rangle\langle\phi^{(2)}|)}_{\mathbf{V}_+} \cdots$$
$$\cdots \underbrace{\left(\frac{1}{z - E^{(r-1)}}|\phi^{(r-1)}\rangle\langle\phi^{(r-1)}|\right)}_{\mathbf{G}_+} \cdot \underbrace{(\mathbf{B}^{(r)} \otimes |\phi^{(r-1)}\rangle\langle\phi^{(r)}|)}_{\mathbf{V}_{+-}}. \tag{4.31}$$

With the notation introduced in Equation 4.31 we could build up an expression for $\mathbf{T}_r$ term by term. As a start, we could express $\mathbf{V}_{-+}$, $\mathbf{V}_{+-}$, $\mathbf{V}_+$, and $\mathbf{G}_+$ as

$$\mathbf{V}_{-+} = \sum_{|\phi\rangle\in\mathcal{L}_-} \sum_{|\phi'\rangle\in\mathcal{L}_+} \mathbf{B}_{\phi,\phi'} \otimes |\phi\rangle\langle\phi'|, \qquad \mathbf{V}_+ = \sum_{|\phi\rangle\in\mathcal{L}_+} \sum_{|\phi'\rangle\mathcal{L}_+} \mathbf{B}_{\phi,\phi'} \otimes |\phi\rangle\langle\phi'|$$
$$\mathbf{V}_{+-} = \sum_{|\phi\rangle\in\mathcal{L}_+} \sum_{|\phi'\rangle\in\mathcal{L}_-} \mathbf{B}_{\phi,\phi'} \otimes |\phi^{(r-1)}\rangle\langle\phi^{(r)}|, \quad \mathbf{G}_+(z) = \sum_{|\phi\rangle\in\mathcal{L}_+} \frac{1}{z - E_\phi}|\phi\rangle\langle\phi|. \tag{4.32}$$

where $\mathbf{B}_{\phi,\phi'}$ is the $\mathbf{B}^{(i)}_{pq,jk}$ block in $\mathbf{V}$ (Equation 4.12 and 4.21) that corresponds to transition from $|\phi\rangle$ to $|\phi'\rangle$, both of which are eigenstates of $\mathbf{H}$. $E_\phi = \langle\phi|H|\phi\rangle$. Multiplying with G_+V_+ gives

$$\begin{aligned}
\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+ &= \sum_{|\phi^{(0)}\rangle\in\mathcal{L}_-} \sum_{|\phi^{(1)}\rangle\in\mathcal{L}_+} \sum_{|\phi^{(2)}\rangle\in\mathcal{L}_+} \left(\mathbf{B}^{(1)} \otimes |\phi^{(0)}\rangle\langle\phi^{(1)}|\right) \cdot \\
&\qquad \cdot \left(\frac{1}{z - E^{(1)}}|\phi^{(1)}\rangle\langle\phi^{(1)}|\right) \left(\mathbf{B}^{(2)} \otimes |\phi^{(1)}\rangle\langle\phi^{(2)}|\right) \\
&= \sum_{|\phi^{(0)}\rangle\in\mathcal{L}_-} \sum_{|\phi^{(1)}\rangle\in\mathcal{L}_+} \sum_{\substack{|\phi^{(2)}\rangle\in\mathcal{L}_+ \\ \|\langle\phi^{(1)}|V|\phi^{(2)}\rangle\|\neq 0}} \mathbf{B}^{(1)} \cdot \frac{1}{z - E^{(1)}} \cdot \mathbf{B}^{(2)} \otimes |\phi^{(0)}\rangle\langle\phi^{(2)}|.
\end{aligned} \tag{4.33}$$

Continue carrying out computations similar in nature to Equation 4.33 to the r^{th} step $|\phi^{(r)}\rangle$ gives us the full expression of $\mathbf{T}_r$ in terms of walks on $\mathbf{H}$ eigenstates in Equation 4.30. ■

We are now ready to derive a general upper bound for $\|\mathbf{T}_r\|_2$ in a similar spirit to Equation 4.20. Following Lemma 4.3.1 as well as properties of ∞-norm mentioned in Section 4.2.3, the 2-norm of $\mathbf{T}_r$ can be bounded from above as

$$\|\mathbf{T}_r\|_2 \leq \max_{|\phi^{(0)}\rangle \in \mathcal{L}_-} \sum_{|\phi^{(r)}\rangle \in \mathcal{L}_-} \left\| \sum\nolimits' \mathbf{B}^{(1)} \cdot \frac{1}{|z - E^{(1)}|} \cdot \mathbf{B}^{(2)} \cdot \frac{1}{|z - E^{(2)}|} \cdots \right.$$
$$\left. \cdots \frac{1}{|z - E^{(r-1)}|} \cdot \mathbf{B}^{(r)} \right\|_\infty \tag{4.34}$$

where the maximum and the first summation are taken over eigenstates of $\mathbf{H}$ in $\mathcal{L}_-$.

Equation 4.34 serves as a starting point for finding tight upper bounds for $\|\mathbf{T}_r\|_2$, because each $\|\mathbf{B}^{(i)}\|_\infty$ can be bounded from above by an appropriate choice of element from the vector $\boldsymbol{\lambda}$ (Definition 4.3.2). In Appendix E we show a concrete example where the upper bound in Equation 4.34 is derived explicitly in terms of elements in $\boldsymbol{\lambda}$ and $\mathbf{M}$.

Since the dimension of the Hilbert space $\mathcal{H}$ grows exponentially as m grows, any algorithm that naively computes the right hand side of Equation 4.34 term by term is likely going to cost $O((D\ell)^{mr})$ where $D = \max_{i=1,\cdots,m} \dim(\mathcal{P}_i)$ is the maximum degeneracy of any subspace. As a first simplification, we could reduce this to $O(\ell^{mr})$ by considering walking in the space of energy configuration (Definition 4.3.4) instead of $\mathbf{H}$ eigenstates.

4.3.4 Walking in the configuration space

The summation in Equation 4.34 is over r-step walks on the $\mathbf{H}$ eigenstates. Note from Equation 4.26 that we could partition eigenstates of $\mathbf{H}$ according to their energy configurations (Definition 4.3.4). We could use this partition simplify this summation by first grouping walks that go through the same changes in energy configurations.

Let $\mathbf{c}^{(i)}$ be the energy configuration of $|\phi^{(i)}\rangle$ in an r-step walk in the space of $\mathbf{H}$ eigenstates. Then the type of walks that appear in terms of $\mathbf{T}_r$ must consist of r steps and satisfy (refer to Definition 4.3.5 for $\mathbf{n}(\mathbf{c})$)

$$\begin{aligned} &\mathbf{n}(\mathbf{c}^{(0)}) \in \mathcal{N}_-, \quad \mathbf{n}(\mathbf{c}^{(r)}) \in \mathcal{N}_- \\ &\mathbf{n}(\mathbf{c}^{(i)}) \in \mathcal{N}_+, \quad i = 1, \cdots, r-1. \end{aligned} \tag{4.35}$$

In other words, the type of walks, or sequences of transitions, must start and end in the low-energy subspace $\mathcal{L}_-$, but stays in the high energy subspace $\mathcal{L}_+$ in between.

Since each term in $\mathbf{V}$ acts on one unperturbed subsystem $\mathcal{H}_i$, at each step which corresponds to the outer product $|\psi^{(i)}\rangle\langle\psi^{(i+1)}|$, the energy configurations $\mathbf{c}^{(i)}$ and $\mathbf{c}^{(i+1)}$ must differ in at most one element. Furthermore, because $\mathbf{V}$ is block-tridiagonal with respect to any subsystem (Equation 4.12), the difference between the respective elements in $\mathbf{c}^{(i)}$ and $\mathbf{c}^{(i+1)}$ must be at most 1. Hence the properties of sequences can be summarized as the following definition.

Definition 4.3.7 (Walk in the configuration space) *We define an* r-step walk in the space of configurations $\mathbf{c}$ *(or* walk in $\mathbf{c}$ for short*) as a sequence of configurations* $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ *such that in addition to satisfying Equation 4.35,* $\{\mathbf{c}^{(i)}\}_{i=0}^{r}$ *also satisfies the property that for every step from* $\mathbf{c}^{(i)}$ *to* $\mathbf{c}^{(i+1)}$ *with* $i = 2, \cdots, r-1$*, either one of the following is true:*

1. $\mathbf{c}^{(i)} = \mathbf{c}^{(i+1)}$*, OR*
2. $\mathbf{c}^{(i+1)}$ *is obtained by incrementing or decrementing one element in* $\mathbf{c}^{(i)}$ *by 1.*

The initial step $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)}$ *and the final step* $\mathbf{c}^{(r-1)} \to \mathbf{c}^{(r)}$ *only satisfy case 2 above.*

The following lemma relates the set of r-step walks in the space of configuration, as defined above, to that in the space of $\mathbf{H}$ eigenstates, as in Definition 4.3.6.

Lemma 4.3.2 *For any* r*-step walk* $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle \to \cdots \to |\phi^{(r)}\rangle$ *described in Definition 4.3.6 there is a walk* $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ *as defined in Definition 4.3.7 such that* $\mathbf{c}(|\phi^{(i)}\rangle) = \mathbf{c}^{(i)}$.

Proof By Definition 4.3.6, $\|\langle\phi^{(i)}|\mathbf{V}|\phi^{(i+1)}\rangle\| \neq 0$ for any $i = 0, \cdots, r-1$. Because of the block tridiagonal structure of $\mathbf{V}$ as in Equation 4.12, the energy configurations $\mathbf{c}(|\phi^{(i)}\rangle)$ and $\mathbf{c}(|\phi^{(i+1)}\rangle)$ differ at at most one element and the difference is at most 1. In particular, the initial step of the walk from $|\phi^{(0)}\rangle \in \mathcal{L}_-$ to $|\phi^{(1)}\rangle \in \mathcal{L}_+$ and the final step from $|\phi^{(r-1)}\rangle \in \mathcal{L}_+$ to $|\phi^{(r)}\rangle \in \mathcal{L}_-$ satisfies $\mathbf{c}(|\phi^{(i)}\rangle) \neq \mathbf{c}(|\phi^{(i+1)}\rangle)$, which fall into case 2 of Definition 4.3.7. Hence if we let $\mathbf{c}^{(i)} = \mathbf{c}(|\phi^{(i)}\rangle)$, the walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ satisfies Definition 4.3.7. ■

For computing a tight upper bound to the ∞-norm of a term in $\|\mathbf{T}_r\|_\infty$ that corresponds to a particular walk satisfying the above Definition 4.3.7, the definitions of λ_i and M_{jk} then come into play. Generally speaking, every step from $\mathbf{c}^{(i)}$ to $\mathbf{c}^{(i+1)}$ contributes a factor. The product of these factors form an upper bound to a term in $\mathbf{T}_r$ that corresponds to an entire walk. If a step falls into the case 1 in the above Definition 4.3.7, then this step contributes a factor ω (Definition 4.3.1). Otherwise if a step falls in the case 2 in Definition 4.3.7 then there must be some element, say the j^{th} element, of $\mathbf{c}_i$ that is changed by 1 to yield the new energy configuration $\mathbf{c}_{i+1}$. The contribution of such a step is λ_j. In other words, a transition has occurred in the subsystem $\mathcal{H}_j$ under the action of $\mathbf{V}$. Further, let j and k be such that the step from $|\psi^{(i)}\rangle$ to $|\psi^{(i+1)}\rangle$ is from the subspace $\mathcal{P}_j$ to $\mathcal{P}_k$ for some subsystem. Then the contributing factor of the step is further multiplied by M_{jk}. To make the above intuition precise, we state the following lemma.

Lemma 4.3.3 *Let f be a function of two energy configurations $\mathbf{c}$ and $\mathbf{c}'$ such that*

$$f(\mathbf{c}, \mathbf{c}') = \begin{cases} \lambda_t M_{ss'} & \mathbf{c} \text{ and } \mathbf{c}' \text{ differ at subsystem } t \text{ where } \mathbf{c}_t = s \text{ and } \mathbf{c}'_t = s' \\ \omega & \mathbf{c} = \mathbf{c}'. \end{cases} \tag{4.36}$$

Then for any $r \geq 3$,

$$\|T_r\|_2 \leq \max_{\mathbf{c}^{(0)}:\mathbf{n}(\mathbf{c}^{(0)})\in\mathcal{N}_-} \sum_{\mathbf{c}^{(r)}:\mathbf{n}(\mathbf{c}^{(r)})\in\mathcal{N}_-} {\sum}'' f(\mathbf{c}^{(0)}, \mathbf{c}^{(1)}) \cdot \frac{1}{|z - E^{(1)}|} \cdots$$

$$\cdots f(\mathbf{c}^{(r-2)}, \mathbf{c}^{(r-1)}) \cdot \frac{1}{|z - E^{(r-1)}|} \cdot f(\mathbf{c}^{(r-1)}, \mathbf{c}^{(r)}). \tag{4.37}$$

Here the summation Σ'' is over all r-step walks in the space of configurations, as defined in Definition 4.3.7, with fixed *initial configuration $\mathbf{c}^{(0)}$. $E^{(i)}$ is the energy of the configuration $\mathbf{c}^{(i)}$, namely $\sum_{j=1}^{m} E_{c_j^{(i)}}$.*

Proof We start from Equation 4.34 and partition the max and summation operations over $\mathbf{H}$ eigenstates according to their energy configurations. Using Lemma 4.3.2 we could deduce from Equation 4.34 that

$$\begin{aligned} \|T_r\|_2 \leq \max_{\mathbf{c}^{(0)}:\mathbf{n}(\mathbf{c}^{(0)})\in\mathcal{N}_-} \max_{|\phi^{(0)}\rangle:\mathbf{c}(|\phi^{(0)}\rangle)=\mathbf{c}^{(0)}} \sum_{\mathbf{c}^{(r)}:\mathbf{n}(\mathbf{c}^{(r)})\in\mathcal{N}_-} {\sum}'' \sum_{\substack{|\phi^{(1)}\rangle\to\cdots\to|\phi^{(r)}\rangle \\ \mathbf{c}(|\phi^{(i)}\rangle)=\mathbf{c}^{(i)}}} \\ \left\| \mathbf{B}^{(1)} \cdot \frac{1}{|z - E^{(1)}|} \cdot \mathbf{B}^{(2)} \cdot \frac{1}{|z - E^{(2)}|} \cdots \frac{1}{|z - E^{(r-1)}|} \cdot \mathbf{B}^{(r)} \right\|_\infty , \end{aligned} \tag{4.38}$$

where the summation Σ'' is defined in the same way as in Equation 4.37. The first two max operations are equivalent to the max operation on the right hand side of Equation 4.34. The three summations essentially sums over the set of all r-step walks on $\mathbf{H}$ eigenstates that are consistent with r-step walks in the space of energy configurations. This set should contain the set of all r-step walks on $\mathbf{H}$ eigenstates that yield non-zero contributions on the right hand side of Equation 4.34. Hence the right hand side of Equation 4.38 is a valid upper bound to that of Equation 4.34. If we remove the max and summation operations over energy configurations in Equation 4.38 by considering a *fixed* walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$, we are left with a term that is bounded from above by

$$\begin{aligned} \max_{|\phi^{(0)}\rangle:\mathbf{c}(|\phi^{(0)}\rangle)=\mathbf{c}^{(0)}} \sum_{\substack{|\phi^{(1)}\rangle\to\cdots\to|\phi^{(r)}\rangle \\ \mathbf{c}(|\phi^{(i)}\rangle)=\mathbf{c}^{(i)}}} \left\|\mathbf{B}^{(1)}\right\|_\infty \cdot \left\| \frac{1}{|z - E^{(1)}|} \cdot \mathbf{B}^{(2)} \cdot \frac{1}{|z - E^{(2)}|} \cdots \right. \\ \left. \cdots \frac{1}{|z - E^{(r-1)}|} \cdot \mathbf{B}^{(r)} \right\|_\infty . \end{aligned} \tag{4.39}$$

Recall that the operator $\mathbf{B}^{(1)}$ is associated with the transition $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle$. The corresponding change in energy configuration is $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)}$. It is established in Lemma

4.3.1 as well as Definition 4.3.7 that $\mathbf{c}^{(0)}$ and $\mathbf{c}^{(1)}$ must differ at one element by 1. Let this be the t^{th} element. In other words, $c_t^{(0)} \neq c_t^{(1)}$. Let $c_t^{(0)} = s$ and $c_t^{(1)} = s'$. We could then interpret $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)}$ as the physical process of a transition in subsystem t from s^{th} energy level to the s'^{th}. Furthermore, $\mathbf{B}^{(1)}$ is the operator associated with transitioning from a *specific* eigenstate $|\phi^{(0)}\rangle$ that satisfies $\langle\phi^{(0)}|\mathbf{H}^{(t)}|\phi^{(0)}\rangle = E_s$, to another $\mathbf{H}$ eigenstate $|\phi^{(1)}\rangle$ with $\langle\phi^{(1)}|\mathbf{H}^{(t)}|\phi^{(1)}\rangle = E_{s'}$. Recall that the superscript (t) for $\mathbf{H}^{(t)}$ represents the t^{th} subsystem, while the superscript for $|\phi^{(i)}\rangle$ stands for the i^{th} step during the walk. Now we are considering all such transitions from $|\phi^{(0)}\rangle$ to $|\phi^{(i)}\rangle$, summing over all possible $|\phi^{(1)}\rangle$ eigenstates and maximizing over all possible $|\phi^{(0)}\rangle$ eigenstates that are consistent with the (fixed) walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots$. By Definition 4.3.2, $\|\mathbf{B}^{(1)}\|_\infty \leq \lambda_t$ for any specific step $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle$. By Definition 4.3.3, there are at most $M_{ss'}$ ways to make a transition from $\mathcal{P}_s$ to $\mathcal{P}_{s'}$ for any subsystem. Hence the contribution of the first step $|\phi^{(0)}\rangle \to |\phi^{(1)}\rangle$ to the right hand side of Expression 4.39 is bounded from above by $\lambda_t M_{ss'}$. Hence Expression 4.39 is bounded from above by

$$f(\mathbf{c}^{(0)}, \mathbf{c}^{(1)}) \cdot \frac{1}{|z - E^{(1)}|} \cdot \max_{|\phi^{(0)}\rangle : \mathbf{c}(|\phi^{(0)}\rangle) = \mathbf{c}^{(0)}} \sum_{\substack{|\phi^{(1)}\rangle \to \cdots \to |\phi^{(r)}\rangle \\ \mathbf{c}(|\phi^{(i)}\rangle) = \mathbf{c}^{(i)}}} \left\|\mathbf{B}^{(2)}\right\|_\infty \cdot \left\| \frac{1}{|z - E^{(2)}|} \cdots \right.$$
$$\left. \cdots \frac{1}{|z - E^{(r-1)}|} \cdot \mathbf{B}^{(r)} \right\|_\infty \tag{4.40}$$

where $f(\mathbf{c}^{(0)}, \mathbf{c}^{(1)}) = \lambda_t M_{ss'}$ following the definition of f in the statement of the Lemma.

The scalar factors $\frac{1}{z - E^{(i)}}$ are constants for all the walks $|\phi^{(1)}\rangle \to \cdots \to |\phi^{(r)}\rangle$ summed over since the walk in configuration space $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ is fixed for Expression 4.39. In other words $E^{(i)} = E(\mathbf{n}(\mathbf{c}^{(i)}))$. The contribution of $\|\mathbf{B}^{(2)}\|_\infty$ could be bounded from above by similar arguments that follow Expression 4.39 that treat $\|\mathbf{B}^{(1)}\|_\infty$, except that one has to consider an alternative possibility when $\mathbf{c}^{(1)} = \mathbf{c}^{(2)}$, in which case the contribution of $\|\mathbf{B}^{(2)}\|_\infty$ over all possible walks on

$\mathbf{H}$ eigenstates is bounded from above by ω (Definition 4.3.1). We could thus bound Expression 4.40 from above by

$$f(\mathbf{c}^{(0)},\mathbf{c}^{(1)})\cdot\frac{1}{|z-E^{(1)}|}\cdot f(\mathbf{c}^{(1)},\mathbf{c}^{(2)})\cdot\frac{1}{|z-E^{(2)}|}\max_{|\phi^{(0)}\rangle:\mathbf{c}(|\phi^{(0)}\rangle)=\mathbf{c}^{(0)}}\sum_{\substack{|\phi^{(1)}\rangle\to\cdots\to|\phi^{(r)}\rangle\\ \mathbf{c}(|\phi^{(i)}\rangle)=\mathbf{c}^{(i)}}}\left\|\mathbf{B}^{(3)}\cdots\frac{1}{|z-E^{(r-1)}|}\cdot\mathbf{B}^{(r)}\right\|_\infty. \tag{4.41}$$

By repeating the arguments that produced Equation 4.41 from Equation 4.40 on $\|\mathbf{B}^{(i)}\|_\infty$ for $i=3,\cdots,r-1$, one could yield upper bounds that are functions of ω, $\boldsymbol{\lambda}$ and $\mathbf{M}$. Finally, apply the same argument for treating $\|\mathbf{B}^{(1)}\|_\infty$ in Expression 4.39 for $\|\mathbf{B}^{(r)}\|_\infty$ yields Equation 4.37. ■

With Lemma 4.3.3 we in essence have accomplished a reduction of the number of walks that need to be enumerated, from $O((D\ell)^{mr})$ as in the case with walks on $\mathbf{H}$ eigenstates in Section 4.3.3, to $O(\ell^{mr})$. In the next section we show how to use symmetry to reduce the exponential dependence on the number of unperturbed subsystems m to polynomial, assuming that both ℓ and r are constant.

4.3.5 Introducing symmetry

In order to further reduce the dimension of the space in which a walk is described, we introduce a symmetric version of the energy configuration. We start by laying down the following definition concerning the status of individual elements in an energy configuration during a walk in the space of $\mathbf{c}$.

Definition 4.3.8 (Active and inactive elements) *Consider an energy configuration $\mathbf{c}^{(i)}$ during a walk in the configuration space $\mathbf{c}^{(0)}\to\cdots\to\mathbf{c}^{(i-1)}\to\mathbf{c}^{(i)}$ with $\mathbf{c}^{(1)}=(0,0,\cdots,0)$. For any k, if the k^{th} element of $c^{(j)}$, which we denote as $c_k^{(j)}$, is 0 for every $j\le i$, then we call $c_k^{(j)}$ an* inactive element. *Otherwise the k^{th} element is an* active element.

In other words, if the k^{th} subsystem is never excited from $\mathcal{P}_0$ during the walk then it is inactive. It is worth noting that an active element of an energy configuration may also be 0. In this case the subsystem was excited from $\mathcal{P}_0$ at some point but returns to $\mathcal{P}_0$.

Definition 4.3.9 (Reduced energy configuration) *For an energy configuration* $\mathbf{c}$ *(Definition 4.3.4) we define* reduced energy configuration $\tilde{\mathbf{c}}$ *as the resulting vector of removing all inactive elements in* $\mathbf{c}$ *and then sorting the active elements in non-decreasing order. In particular, let* $\tilde{\mathbf{c}}(\mathbf{c})$ *be the reduced energy configuration that corresponds to a configuration* $\mathbf{c}$*.*

For example, in a setting with $m = 3$ subsystems, the configuration where the first subsystem has energy E_3, the second is inactive and thus has energy E_0, the third has E_1 and the fourth has E_0 but is active would have an energy configuration $\mathbf{c} = (3, 0, 1, 0)$. However, in this case the reduced energy configuration $\tilde{\mathbf{c}} = (0, 1, 3)$. If the second subsystem is active then $\tilde{\mathbf{c}} = (0, 0, 1, 3)$ is the reduced energy configuration.

The advantage of introducing this concept is that the space in which the walks are described can be reduced from exponential in m to polynomial, assuming both ℓ, the total number of energy levels in each unperturbed subsystem, and r, the order of the perturbation or the total number of steps in a walk, are constant. For a fixed set of parameters m, ℓ, the total possible energy configurations $\mathbf{c}$ is $O(\ell^m)$. However, as we show in the following lemma, the set of a possible reduced energy configuration $\tilde{\mathbf{c}}$ is polynomial in m.

Lemma 4.3.4 *Let* $f_{m\ell}$ *be the total number of possible reduced energy configurations of length* m *and maximum possible number of energy levels* ℓ*. Then* $f_{m\ell} \leq m^\ell$ *for any* $m \geq 2$ *and* $\ell \geq 1$*.*

Proof The last element of a reduced configuration could take any one of ℓ values. Since by Definition 4.3.9, the elements of a reduced configuration is non-decreasing, the remaining $m-1$ elements of $\tilde{\mathbf{c}}$ has $f_{m-1,\tilde{\mathbf{c}}_m}$ choices where $\tilde{\mathbf{c}}_m \in \{0, \cdots, \ell-1\}$ is the

last element of $\tilde{\mathbf{c}}$. We then have the recursion $f_{m\ell} = f_{m-1,\ell} + f_{m-1,\ell-1} + \cdots + f_{m-1,1}$ with boundary condition $f_{k1} = 1$ for any $k \in \{1, \cdots, m\}$ and $f_{1k} = k$ for any $k \in \{0, 1, \cdots, \ell - 1\}$. Hence $f_{m\ell} = f_{m-1,\ell} + f_{m,\ell-1} = 1 + \sum_{i=1}^{m} f_{i,\ell-1}$. Starting from $f_{m1} = 1$, we have $f_{m2} = 1 + f_{11} + f_{21} + \cdots + f_{m1} \leq 1 + mf_{m1} = 1 + m$ and $f_{m3} = 1 + f_{12} + f_{22} + \cdots + f_{m2} \leq 1 + m + m^2$. Applying this to $f_{m\ell}$, we have $f_{m\ell} \leq 1 + f_{m,\ell-1} \leq 1 + m(1 + mf_{m,\ell-2}) \leq \cdots \leq 1 + m + \cdots + m^{\ell-1} \leq m^{\ell}$. ■

We now define the notion of walks in the reduced configuration space as the follows.

Definition 4.3.10 (Walk in the space of reduced configurations) *A sequence of reduced configurations* $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \to \tilde{\mathbf{c}}^{(r-1)} \to \tilde{\mathbf{c}}^{(r)}$ *is an* r-step walk in the space of reduced configurations $\tilde{\mathbf{c}}$ *if*

$$\begin{aligned} &\mathbf{n}(\tilde{\mathbf{c}}^{(0)}) = \mathbf{n}(\tilde{\mathbf{c}}_0) \in \mathcal{N}_-, \quad \mathbf{n}(\tilde{\mathbf{c}}^{(r)}) \in \mathcal{N}_- \\ &\mathbf{n}(\tilde{\mathbf{c}}^{(i)}) \in \mathcal{N}_+, \quad i = 1, \cdots, r-1. \end{aligned} \tag{4.42}$$

and either one of the following is true for any $i = 2, \cdots, r-1$*:*

1. $\tilde{\mathbf{c}}^{(i)} = \tilde{\mathbf{c}}^{(i+1)}$*, OR*
2. $\tilde{\mathbf{c}}^{(i)}$ *and* $\tilde{\mathbf{c}}^{(i+1)}$ *differ by 1 at one element, OR*
3. $|\tilde{\mathbf{c}}^{(i+1)}| = |\tilde{\mathbf{c}}^{(i)}| + 1$.

As a consequence, for the initial step $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)}$ *only case 3 applies and for the final step* $\tilde{\mathbf{c}}^{(r-1)} \to \tilde{\mathbf{c}}^{(r)}$ *only case 2 applies.*

The following lemma connects the space of reduced energy configurations $\tilde{\mathbf{c}}$ to that of energy configuration $\mathbf{c}$.

Lemma 4.3.5 *For every walk* $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ *in the space of* $\mathbf{c}$ *as in Definition 4.3.7, there is a corresponding walk* $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \to \tilde{\mathbf{c}}^{(r)}$ *in the space of* $\tilde{\mathbf{c}}$ *as in Definition 4.3.10 such that* $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}^{(i)}$*. Furthermore, for any permutation*

π over m elements, the walk $\pi(\mathbf{c}^{(0)}) \to \pi(\mathbf{c}^{(1)}) \to \cdots \to \pi(\mathbf{c}^{(r)})$ also maps to the same walk in $\tilde{\mathbf{c}}$. Conversely, for any walk $\mathbf{c}'^{(0)} \to \mathbf{c}'^{(1)} \to \cdots \to \mathbf{c}'^{(r)}$ that satisfies both Definition 4.3.7 and $\tilde{\mathbf{c}}(\mathbf{c}'^{(i)}) = \tilde{\mathbf{c}}^{(i)}$, there must be a permutation π' such that $\pi'(\mathbf{c}^{(i)}) = \mathbf{c}'^{(i)}$ for any i.

Proof By definition, $\mathbf{n}(\mathbf{c}^{(0)}) \in \mathcal{N}_-$. Since the definition of energy combination $\mathbf{n}$ (Definition 4.3.5) is invariant with respect to permutation of unperturbed subsystems, $\mathbf{n}(\tilde{\mathbf{c}}(\mathbf{c}^{(0)})) = \mathbf{n}(\mathbf{c}^{(0)}) \in \mathcal{N}_-$. For every subsequent step $\mathbf{c}^{(i)} \to \mathbf{c}^{(i+1)}$, $i \in \{0, \cdots, r-2\}$, case 1 in Definition 4.3.7 leads to $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}(\mathbf{c}^{(i+1)})$, which fits case 1 of Definition 4.3.10. Case 2 in Definition 4.3.7 depends on whether an inactive element in $\mathbf{c}^{(i)}$ becomes active in $\mathbf{c}^{(i+1)}$. If this is not the case, then $\tilde{\mathbf{c}}(\mathbf{c}^{(i)})$ and $\tilde{\mathbf{c}}(\mathbf{c}^{(i+1)})$ differ by 1 at one element, matching case 2 in Definition 4.3.10. Otherwise the additional active element in $\mathbf{c}^{(i+1)}$ contributes an additional element in $\tilde{\mathbf{c}}(\mathbf{c}^{(i+1)})$, namely $|\tilde{\mathbf{c}}(\mathbf{c}^{(i+1)})| = |\tilde{\mathbf{c}}(\mathbf{c}^{(i)})| + 1$. Finally from $\mathbf{n}(\mathbf{c}^{(r)}) \in \mathcal{N}_-$ we have $\mathbf{n}(\tilde{\mathbf{c}}(\mathbf{c}^{(r)})) \in \mathcal{N}_-$. Hence if we let $\tilde{\mathbf{c}}^{(i)} = \tilde{\mathbf{c}}(\mathbf{c}^{(i)})$ then the walk $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \to \tilde{\mathbf{c}}^{(r)}$ matches the Definition 4.3.10. This proves the first part of the lemma.

The second part follows by noting that by Definition 4.3.9, the reduced energy configuration of an $\mathbf{H}$ eigenstate is invariant with respect to permutation of the subsystems, namely $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}(\pi(\mathbf{c}^{(i)}))$ for any permutation π over m elements.

The last part ("Conversely...") can be proved by starting with the observation that for any walk $\mathbf{c}'^{(0)} \to \mathbf{c}'^{(1)} \to \cdots \to \mathbf{c}'^{(r)}$ that satisfies both Definition 4.3.10 and $\tilde{\mathbf{c}}(\mathbf{c}'^{(i)}) = \tilde{\mathbf{c}}^{(i)}$, because $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}^{(i)}$ and by the permutation invariance of reduced energy configuration there must be a permutation $\pi^{(i)}$ such that $\pi^{(i)}(\mathbf{c}^{(i)}) = \mathbf{c}'^{(i)}$ for every $i \in \{0, \cdots, r\}$. Our goal is thus to show that the permutations $\pi^{(i)}$ are identical to the same permutation π'. For the sake of contradiction suppose $\pi^{(i)} \neq \pi^{(i+1)}$ for some i. Then there must be a (non-trivial) permutation $\Delta\pi$ such that $\pi^{(i+1)} = \Delta\pi \cdot \pi^{(i)}$. Since the walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ conforms to Definition 4.3.7, either $\mathbf{c}^{(i)} = \mathbf{c}^{(i+1)}$ or $\mathbf{c}^{(i)}$ and $\mathbf{c}^{(i+1)}$ differ by 1 at one element. We discuss each case individually as the following:

- Suppose $\mathbf{c}^{(i)} = \mathbf{c}^{(i+1)}$, then $\mathbf{c}'^{(i)} = \pi^{(i)}(\mathbf{c}^{(i)}) = \pi^{(i)}(\mathbf{c}^{(i+1)})$. Hence $\mathbf{c}'^{(i+1)} = \pi^{(i+1)}(\mathbf{c}^{(i+1)}) = \Delta\pi(\pi^{(i)}(\mathbf{c}^{(i+1)})) = \Delta\pi(\mathbf{c}'^{(i)})$, which is impossible if the walk $\mathbf{c}'^{(0)} \to \mathbf{c}'^{(1)} \to \cdots \to \mathbf{c}'^{(r)}$ conforms to Definition 4.3.7 because no step $\mathbf{c}'^{(i)} \to \mathbf{c}'^{(i+1)}$ that conforms to case 1 or 2 in Definition 4.3.7 corresponds to a non-trivial permutation of $\mathbf{c}'^{(i)}$. Hence in this case the permutations $\pi^{(i)}$ and $\pi^{(i+1)}$ must be identical.

- Suppose $\mathbf{c}^{(i)}$ and $\mathbf{c}^{(i+1)}$ differ by 1 at one element, namely $\mathbf{c}_j^{(i)} \neq \mathbf{c}_j^{(i+1)}$ for some j. Then $\pi^{(i)}(\mathbf{c}^{(i)})$ and $\pi^{(i)}(\mathbf{c}^{(i+1)})$ differ at an element $k \neq j$. Since $\mathbf{c}'^{(i+1)} = \Delta\pi(\pi^{(i)}(\mathbf{c}^{(i+1)}))$ and $\mathbf{c}'^{(i)} = \pi^{(i)}(\mathbf{c}^{(i)})$, we see that the step $\mathbf{c}'^{(i)} \to \mathbf{c}'^{(i+1)}$ is realized by incrementing the k^{th} element of $\mathbf{c}'^{(i)}$ by $\mathbf{c}_j^{(i+1)} - \mathbf{c}_j^{(i)}$ and apply a non-trivial permutation $\Delta\pi$. The latter step contradicts Definition 4.3.7 since no permutation is possible in a single step with either case 1 or 2 in Definition 4.3.7.

Therefore we have shown that the set of r-step walks in $\mathbf{c}$ that is consistent with a particular r-step walk in $\tilde{\mathbf{c}}$ are merely the same walk in $\mathbf{c}$ with different permutations of the unperturbed subsystems. ■

We could then establish an upper bound for $\|\mathbf{T}_r\|_2$ that is based on a walk in the space of $\tilde{\mathbf{c}}$ as in Definition 4.3.10, which is stated in the following Lemma.

Lemma 4.3.6 *For an r-step walk $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \to \tilde{\mathbf{c}}^{(r)}$ in the space of reduced configuration $\tilde{\mathbf{c}}$ as described in Definition 4.3.10, consider any r-step walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ such that $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}^{(i)}$. Define the set $\mathcal{F}_i = \{j = 1, \cdots, r | c_i^{(j-1)} \neq c_i^{(j)}\}$, the vector $\mathbf{f} \in \mathbb{N}^m$ such that $\mathbf{f}_i = |\mathcal{F}_i|$ and an integer $k = r - \sum_{i=1}^m \mathbf{f}_i$. Let $\mathbf{b} \in \mathbb{N}^m$ be $\mathbf{f}$ sorted in non-increasing order (to match Definition 4.2.2). Then*

$$\|\mathbf{T}_r\|_2 \leq \max_{\tilde{\mathbf{c}}^{(0)}:\mathbf{n}(\tilde{\mathbf{c}}^{(0)})\in\mathcal{N}_-} \sum_{\tilde{\mathbf{c}}^{(r)}:\mathbf{n}(\tilde{\mathbf{c}}^{(r)})\in\mathcal{N}_-} \sum^* \left(\prod_{i=1}^r \frac{1}{|z - E^{(i)}|}\right) \cdot m_{\mathbf{b}}(\boldsymbol{\lambda}) \cdot \left(\prod_{j:\exists i, j\in\mathcal{F}_i} M_{\mathbf{c}_i^{(j-1)}, \mathbf{c}_i^{(j)}}\right) \cdot \omega^k \tag{4.43}$$

where ω, $\boldsymbol{\lambda}$ and $\mathbf{M}$ are defined in Definitions 4.3.1, 4.3.2 and 4.3.3 respectively. The summation Σ^* is over all r-step walks in the space of reduced configurations, as defined in Definition 4.3.10, with fixed *initial reduced configuration $\tilde{\mathbf{c}}^{(0)}$ and final reduced configuration $\tilde{\mathbf{c}}^{(r)}$.*

Proof Starting from Lemma 4.3.3, where we bounded from above contributions of individual r-step walks in $\mathbf{c}$ by an expression

$$f(\mathbf{c}^{(0)},\mathbf{c}^{(1)})\cdot\frac{1}{|z-E^{(1)}|}\cdot f(\mathbf{c}^{(1)},\mathbf{c}^{(2)})\cdots f(\mathbf{c}^{(r-2)},\mathbf{c}^{(r-1)})\cdot\frac{1}{|z-E^{(r-1)}|}\cdot f(\mathbf{c}^{(r-1)},\mathbf{c}^{(r)}). \tag{4.44}$$

For a specific r-step walk in $\mathbf{c}$ space, let $\mathcal{F}_i=\{j=1,\cdots,r|c_i^{(j-1)}\neq c_i^{(j)}\}$. Then using the definition of $f(\mathbf{c},\mathbf{c}')$ in Lemma 4.3.3, we could rewrite expression 4.44 as

$$\left(\prod_{i=1}^{m}\frac{1}{|z-E^{(i)}|}\right)\cdot\left(\prod_{i=1}^{m}\lambda_i^{|\mathcal{F}_i|}\right)\cdot\left(\prod_{j:\exists i,j\in\mathcal{F}_i}M_{\mathbf{c}_i^{(j-1)},\mathbf{c}_i^{(j)}}\right)\cdot\omega^k. \tag{4.45}$$

For a *fixed* walk $\tilde{\mathbf{c}}^{(0)}\to\tilde{\mathbf{c}}^{(1)}\to\cdots\to\tilde{\mathbf{c}}^{(r)}$, consider the set $\mathcal{W}$ of r-step walks in the space of $\mathbf{c}$ such that $\tilde{\mathbf{c}}(\mathbf{c}^{(i)})=\tilde{\mathbf{c}}^{(i)}$. By Lemma 4.3.5, $\mathcal{W}$ consists of permutations of some r-step walk in $\mathbf{c}$. If the contribution of a single walk in $\mathcal{W}$ can be bounded from above by Equation 4.45, then the total contribution from the walks in $\mathcal{W}$ can be bounded from above by summing over *all* possible permutations of the unperturbed subsystems, namely

$$\sum_{\pi:[m]\mapsto[m]}\left(\prod_{i=1}^{m}\frac{1}{|z-E^{(i)}|}\right)\cdot\left(\prod_{i=1}^{m}\lambda_{\pi(i)}^{|\mathcal{F}_i|}\right)\cdot\left(\prod_{j:\exists i,j\in\mathcal{F}_i}M_{\mathbf{c}_i^{(j-1)},\mathbf{c}_i^{(j)}}\right)\cdot\omega^k. \tag{4.46}$$

Because the reduced energy configuration $\tilde{\mathbf{c}}$ is invariant with respect to the energy configuration $\mathbf{c}$ that it corresponds to, we have

$$\prod_{j:\exists i,j\in\mathcal{F}_i}M_{\mathbf{c}_i^{(j-1)},\mathbf{c}_i^{(j)}}=\prod_{j:\exists i,j\in\mathcal{F}_i'}M_{\mathbf{c}_i'^{(j-1)},\mathbf{c}_i'^{(j)}} \tag{4.47}$$

for any $\mathbf{c}'^{(0)}\to\mathbf{c}'^{(1)}\to\cdots\to\mathbf{c}'^{(r)}$ such that $\pi(\mathbf{c}^{(i)})=\mathbf{c}'^{(i)}$ for some permutation π. Here $\mathcal{F}_i'=\{j=1,\cdots,r|c_i'^{(j-1)}\neq c_i'^{(j)}\}$. Then by Definition 4.2.1 and 4.2.2, $\sum_{\pi:[m]\mapsto[m]}\lambda_{\pi(i)}^{\mathbf{b}_i}=m_{\mathbf{b}}(\boldsymbol{\lambda})$ where $\mathbf{b}$ is defined in the statement of the Lemma. Expression 4.46 serves as an upper bound for a *fixed* walk in $\tilde{\mathbf{c}}$. Summing over all r-step

walks in $\tilde{\mathbf{c}}$ described in Definition 4.3.10, and incorporating Equation 4.47, we can bound the right hand side of Equation 4.37 by that of Equation 4.43. ■

In Figure 2 of the main text we have already demonstrated the relationship between a walk in $\mathbf{c}$ and a walk in $\tilde{\mathbf{c}}$. Furthermore, we presented Equation (5) in the main text without proof. In Appendix F we illustrate Lemma 4.3.6 with a concrete derivation of Equation (5) of the main text, in order to provide more intuitive arguments for understanding the construction of the upper bound in Equation 4.43.

4.4 Efficient algorithm for computing upper bounds

4.4.1 Constructing cellular automaton

In Definition 4.3.5 for energy combination, we define $\mathcal{C}(\mathbf{n})$ as the set of energy configurations that give rise to the energy combination $\mathbf{n}$, while $\mathbf{n}(\mathbf{c})$ is the energy combination corresponding to a given energy configuration. Note that the mapping from an energy combination to an energy configuration is not unique (since for example $\mathbf{c} = (0, 1)$ and $\mathbf{c} = (1, 0)$ both correspond to $\mathbf{n} = (1, 1)$) while the mapping in the reverse direction is unique. To enforce uniqueness in both directions, we define *uniquely reduced energy configuration* as the following.

Definition 4.4.1 (Uniquely reduced energy configuration) *Referring to Definition 4.3.4, for an energy configuration* $\mathbf{c}$ *we define* uniquely reduced energy configuration $\hat{\mathbf{c}}$ *as the resulting vector of removing* all *zero elements in* $\mathbf{c}$ *and then sorting the active elements in ascending order. For each energy combination* $\mathbf{n}$ *let* $\hat{\mathbf{c}}(\mathbf{n})$ *be the uniquely reduced energy configuration corresponding to* $\mathbf{n}$.

Note that Definition 4.4.1 is only minutely different from Definition 4.3.9 in terms of which zero elements to remove. With Definition 4.4.1 for each energy combination $\mathbf{n}$ there is a unique $\hat{\mathbf{c}}$ that is consistent with $\mathbf{n}$. For example consider $\mathbf{c}_1 = (0, 1, 0, 3)$ and $\mathbf{c}_2 = (0, 0, 3, 1)$, both of which belong in the set $\mathcal{C}(\mathbf{n})$ with $\mathbf{n} = (1, 0, 1)$, but we

have a unique $\hat{\mathbf{c}} = (1, 3)$ that corresponds to $\mathbf{n} = (1, 0, 1)$. In fact it is not hard to see that

$$\hat{\mathbf{c}}(\mathbf{n}) = (\underbrace{1, \cdots, 1}_{n_1}, \underbrace{2, \cdots, 2}_{n_2}, \cdots, \underbrace{\ell, \cdots, \ell}_{n_\ell}). \tag{4.48}$$

Our cellular automaton then consists of cells (graph nodes) connected with directed edges. Each cell is associated with a list of 4-tuples $(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu)$. An n-tuple is an ordered sequence of n elements. Here in our 4-tuple, $\tilde{\mathbf{c}}$ is a reduced energy configuration (Definition 4.3.9) and $\mathbf{b}$ is a reduced partition vector (Definition 4.2.2), ξ is a scalar coefficient and $\mu : \tilde{\mathbf{c}} \mapsto \mathbf{b}$ is a one-one mapping from the reduced energy configuration to the reduced partition. Because of its bijective nature, one could also think of μ as a permutation map. The reason for introducing the mapping μ is because the reduced partition does not contain all the information about the current configuration.

We construct the cellular automaton with BuildCA subroutine as described in Algorithm 2. The algorithm produces a directed graph $G(\mathcal{V}, \mathcal{E})$ that represents the cellular automaton. Each node $v_{\mathbf{n}} \in \mathcal{V}$ corresponds to an energy combination $\mathbf{n}$. In each node $v_{\mathbf{n}}$ and each directed edge $e(v_{\mathbf{n}}, v_{\mathbf{n}'}) \in \mathcal{E}$ we store a list of 4-tuples $(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu)$ denoted as $\mathcal{S}_{\mathbf{n}}$ and $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$ respectively. For a given energy combination vector $\mathbf{n} = (n_0, n_1, n_2, \cdots, n_{\ell-1})$, we introduce the notation

$$\begin{aligned} \mathbf{n}_0 &= \underbrace{(m, 0, \cdots, 0)}_{\ell} \\ \mathbf{n}'_i &= (n_1, \cdots, n_i - 1, n_{i+1} + 1, \cdots, n_{\ell-1}), \qquad i = 1, \cdots, \ell - 2. \end{aligned} \tag{4.49}$$

For an energy combination $\mathbf{n}$ to be *compatible* with our physical setting (Figure 1 of the main text), it is necessary that

$$\sum_{i=0}^{\ell-1} n_i \leq m, \quad \text{and} \quad n_i \geq 0, \quad n_i \in \mathbb{Z}, \quad \forall i \in \{0, \cdots, \ell - 1\}. \tag{4.50}$$

Note that the definition of $\mathbf{n}'_i$ in Equation 4.49 for a given $\mathbf{n}$ essentially corresponds to a step $\mathbf{c}^{(j)} \to \mathbf{c}^{(j+1)}$ in the space of configurations $\mathbf{c}$ where $\mathbf{c}^{(j+1)}$ and $\mathbf{c}^{(j)}$ differ by

1 at one subsystem and going from $\mathbf{c}^{(j)}$ to $\mathbf{c}^{(j+1)}$ the subsystem makes a transition from energy level i to $i+1$. The graph $G(\mathcal{V}, \mathcal{E})$ that Algorithm 2 connects any energy combination $\mathbf{n}$ with another energy combination $\mathbf{n}'$ as long as there is a walk in $\mathbf{c}$ (Definition 4.3.7) such that at some step j, $\mathbf{n}(\mathbf{c}^{(j)}) = \mathbf{n}$ and $\mathbf{n}(\mathbf{c}^{(j+1)}) = \mathbf{n}'$.

Since the energy combination $\mathbf{n}$ is a vector of length ℓ and each element of $\mathbf{n}$ takes values from $[m]$, there are in total $O(m^{\ell+1})$ possible energy combinations. The most naive implementation of Algorithm 2 takes $O(m^{2(\ell+1)})$. If we consider ℓ to be a constant for the physical system, Algorithm 2 costs computational resource that is polynomial in the system size m.

4.4.2 Cell update rules

Recall that we are interested in computing an upper bound for $\|\mathbf{T}_r\|_\infty$ for any r. The goal of this section is to present the update rules for each individual cells so that in the end the upper bound for $\|\mathbf{T}_r\|_\infty$ can be gleaned from all nodes $v_{\mathbf{n}}$ such that $\mathbf{n} \in \mathcal{N}_-$ after r concurrent updates for all nodes in the cellular automaton.

Let $\mathcal{S}_{\mathbf{n}}$ be the set of 4-tuples associated with the cell $v_{\mathbf{n}}$. To aid the presentation we define a scalar multiplication rule for the 4-tuples: $C(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu) \equiv (\tilde{\mathbf{c}}, \mathbf{b}, C\xi, \mu)$ where C is a scalar quantity. Naturally we extend the multiplication rule to entire sets of the 4-tuples:

$$C\mathcal{S}_{\mathbf{n}} \equiv \{(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu) \in \mathcal{S}_{\mathbf{n}} | (\tilde{\mathbf{c}}, \mathbf{b}, C\xi, \mu)\}.$$

Similarly we define $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$ as the set of 4-tuples associated with the edge $e(\mathbf{n}, \mathbf{n}') \in \mathcal{E}$. The rules for updating $\mathcal{S}_{\mathbf{n}}$ for each cell $v_{\mathbf{n}}$ and $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$ for any edge $e(\mathbf{n}, \mathbf{n}')$ is outlined in the UpdateCell subroutine in Algorithm 3.

The procedure UpdateCell($v_{\mathbf{n}}$) called on a particular cell $v_{\mathbf{n}}$ contains two main steps: the first updates the tuple list $\mathcal{S}_{\mathbf{n}}$ of the current cell by combining $\mathcal{S}_{\mathbf{n}}$ scaled by $\omega/(z - E(\mathbf{n}))$ with the tuple lists on the incident edges scaled by $1/(z - E(\mathbf{n}))$. See Equation 4.51. The second step is to generate 4-tuple lists for the outgoing edges from the current cell by the Out($\mathbf{n}, \mathbf{n}', \mathcal{T}$) subroutine. During the first step, the factor

Algorithm 2: Cellular automaton construction algorithm

Input: The number of subsystems m as shown in Figure 1a of the main text; The matrix $\mathbf{M} \in \mathbb{R}^{\ell \times \ell}$ as in Definition 4.3.3.

Output: A weighted directed graph $G(\mathcal{V}, \mathcal{E})$ that serves as a representation of the cellular automaton.

Procedure $G(\mathcal{V}, \mathcal{E}) = \textsc{BuildCA}(m, \mathbf{M})$

1. $\mathcal{V} \leftarrow \{v_{\mathbf{n}_0}\}$, $\mathcal{E} \leftarrow \emptyset$;
2. $\textsc{BuildCell}(\mathbf{n}_0)$;
3. Return $G(\mathcal{V}, \mathcal{E})$.

Procedure $\textsc{BuildCell}(\mathbf{n})$

1. For each $t = 0, 1, \cdots, \ell - 1$, compute $\mathbf{n}'_t$ and test if it satisfies (4.50). If so, then
   - If $v_{\mathbf{n}'_i} \notin \mathcal{V}$, $\quad \mathcal{V} \leftarrow \mathcal{V} \cup \{v_{\mathbf{n}'_i}\}$;
   - If $e(v_{\mathbf{n}}, v_{\mathbf{n}'_i}) \notin \mathcal{E}$, $\quad \textsc{AddEdge}(\mathbf{n}, \mathbf{n}'_t)$;
     If $e(v_{\mathbf{n}'_i}, v_{\mathbf{n}}) \notin \mathcal{E}$, $\quad \textsc{AddEdge}(\mathbf{n}'_t, \mathbf{n})$;
2. If $\mathbf{n} = (0, \cdots, 0, m)$, return.
   Otherwise for each $t = 0, 1, \cdots, \ell - 1$, call $\textsc{BuildCell}(\mathbf{n}'_t)$.

Procedure $\textsc{AddEdge}(\mathbf{p}, \mathbf{q})$

1. Find s and t such that $q_s = p_s - 1$ and $q_t = p_t + 1$;
2. Add $e(v_{\mathbf{p}}, v_{\mathbf{q}})$ with weight M_{st}.

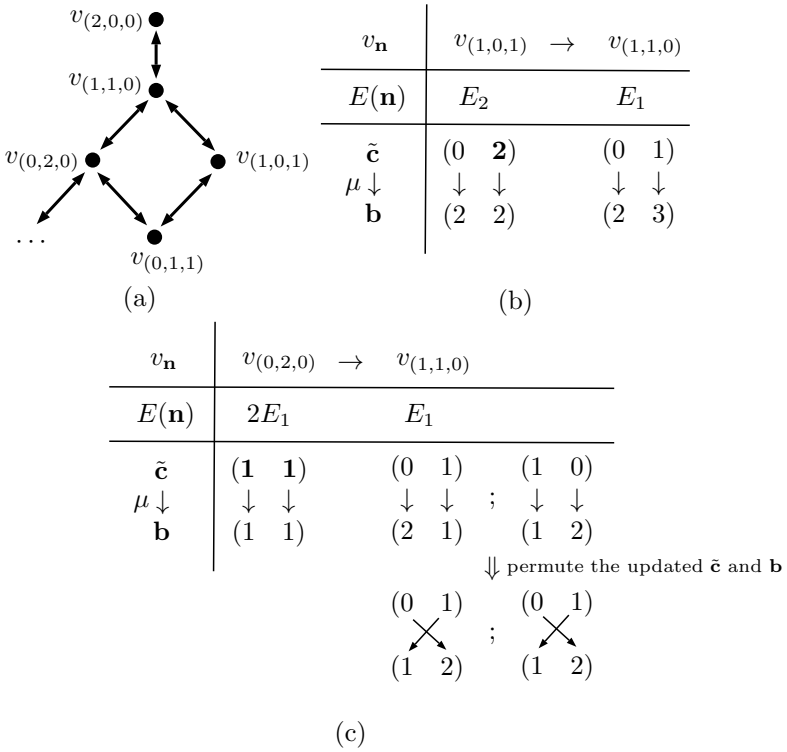


Fig. 4.6. An example illustrating the OUT$(\mathbf{n}, \mathbf{n}', \mathcal{T})$ subroutine in Algorithm 3. Here we let the total number of subsystems be $m = 2$ and each of them has $\ell = 3$ energy levels. (a): The graph $G(\mathcal{V}, \mathcal{E})$ generated by Algorithm 2. Here only part of $\mathbf{G}$ is shown. (b): During a call for OUT$(\mathbf{n}, \mathbf{n}', \mathcal{T})$ with $\mathbf{n} = (0, 1)$ and $\mathbf{n}' = (1, 0)$, the 4-tuple $\mathcal{T} = (\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu) \in \mathcal{S}_{\mathbf{n}}$ with $\tilde{\mathbf{c}} = (0, \mathbf{2})$ and $\mathbf{b} = (2, 2)$, which is shown in the left column of (b), is being used for generating a new 4-tuple $(\tilde{\mathbf{c}}_{\text{new}}, \mathbf{b}_{\text{new}}, \xi_{\text{new}}, \mu_{\text{new}}) \in \mathcal{S}_{\mathbf{n},\mathbf{n}'}$ with $\tilde{\mathbf{c}}_{\text{new}} = (0, 1)$ and $\mathbf{b}_{\text{new}} = (2, 3)$. Here the bold $\mathbf{2}$ in $\tilde{\mathbf{c}}$ represents the "marked" element in step 3b of Algorithm 3. Note that $\mathbf{n}(\mathbf{c}) = \mathbf{n}$ and $\mathbf{n}(\mathbf{c}_{\text{new}}) = \mathbf{n}'$. (c): During a call for OUT$(\mathbf{n}, \mathbf{n}', \mathcal{T})$ with $\mathbf{n} = (1, 1)$ and $\mathbf{n}' = (1, 0)$, similar to (b) we use the 4-tuple $\mathcal{S}_{\mathbf{n}}$ to generate new 4-tuples to be stored in $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$. However, here both elements of $\tilde{\mathbf{c}} = (\mathbf{1}, \mathbf{1})$ are "marked". Hence step 3c of Algorithm 3 generates two new 4-tuples, each with their $\tilde{\mathbf{c}}_{\text{new}}$ having one distinct element that differs its counterpart in $\tilde{\mathbf{c}}$ by 1. The step with the label "permute the updated $\tilde{\mathbf{c}}$ and $\mathbf{b}$" illustrates the step 6 in OUT in Algorithm 3, where elements of $\mathbf{c}_{\text{new}}$ and $\mathbf{b}_{\text{new}}$ as well as the mapping $\mu_{\text{new}} : \mathbf{c}_{\text{new}} \mapsto \mathbf{b}_{\text{new}}$ are arranged to conform to their respective definitions (Definition 4.3.9 for $\tilde{\mathbf{c}}$ and Definition 4.2.2 for $\mathbf{b}$).

$1/(z - E(\mathbf{n}))$ is to account for the contribution of $\mathbf{G}_+$ terms in $\mathbf{T}_r$. The ω factor in the first step is to account for the case where the walk in $\tilde{\mathbf{c}}$ (or $\mathbf{c}$) stays at the same configuration. The second step is to compute the correct list of 4-tuples to deliver to each $\mathbf{n}'$ in the next update. For each $\mathbf{n}'$ that is accessible from the current energy combination $\mathbf{n}$, each 4-tuple in $\mathcal{S}_{\mathbf{n}}$ will contribute an appropriate set of 4-tuples that are stored in $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$. These new 4-tuples must conform to the transition from $\mathbf{n}$ to $\mathbf{n}'$, in the sense that is demonstrated in Figure 4.6. We will make these intuition precise in the next section, where we prove Theorems 4.4.1 and 4.4.2.

Algorithm 3: Updating the cells and their outgoing edges

Input:

- The node $v_{\mathbf{n}} \in \mathcal{V}$ from the graph $G(\mathcal{V}, \mathcal{E})$ with $\mathcal{E} = \mathcal{E}_{\text{dashed}} \cup \mathcal{E}_{\text{non-dashed}}$ generated by Algorithm 2.

Output:

- Updated list of 4-tuples $\mathcal{S}_{\mathbf{n}}$ associated with $v_{\mathbf{n}}$, and $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$ associated with each outgoing edge $e(v_{\mathbf{n}}, v_{\mathbf{n}'}) \in \mathcal{E}$.

Procedure UpdateCell($v_{\mathbf{n}}$)

1. Update the list of 4-tuples at each cell $v_{\mathbf{n}}$:

$$\mathcal{S}_{\mathbf{n}} \leftarrow \left(\frac{\omega}{|z - E(\mathbf{n})|} \mathcal{S}_{\mathbf{n}} \right) \cup \left(\frac{1}{|z - E(\mathbf{n})|} \bigcup_{\mathbf{n}' : e(\mathbf{n}', \mathbf{n}) \in \mathcal{E}} \mathcal{S}_{\mathbf{n}', \mathbf{n}} \right). \tag{4.51}$$

2. For each outgoing edge $e(v_{\mathbf{n}}, v'_{\mathbf{n}}) \in \mathcal{E}$ do the following

$$\mathcal{S}_{\mathbf{n},\mathbf{n}'} \leftarrow \bigcup_{\mathcal{T} \in \mathcal{S}_{\mathbf{n}}} \textsc{Out}(\mathbf{n}, \mathbf{n}', \mathcal{T}) \tag{4.52}$$

   where Out is a subroutine described in the Out subroutine.

Procedure $\mathcal{S}_{\text{new}} = \textsc{Out}(\mathbf{n}, \mathbf{n}', \mathcal{T})$

1. $\xi_{\text{new}} \leftarrow M_{\mathbf{n},\mathbf{n}'}\xi$, where $M_{\mathbf{n},\mathbf{n}'}$ is the weight of the edge $e(v_{\mathbf{n}}, v_{\mathbf{n}'})$.
2. Compute $\tilde{\mathbf{c}} = (\underbrace{0, \cdots, 0}_{|\mathbf{b}|-|\hat{\mathbf{c}}(\mathbf{n})|}, \hat{\mathbf{c}}(\mathbf{n}))$ and $\tilde{\mathbf{c}}' = (\underbrace{0, \cdots, 0}_{|\mathbf{b}|-|\hat{\mathbf{c}}(\mathbf{n})|}, \hat{\mathbf{c}}(\mathbf{n}'))$.
3. If $|\hat{\mathbf{c}}(\mathbf{n})| \geq |\hat{\mathbf{c}}(\mathbf{n}')|$,
   (a) Find k such that $\tilde{c}_k \neq \tilde{c}'_k$ and compute $\Delta c = \tilde{c}'_k - \tilde{c}_k$.
   (b) Mark all $\tilde{c}_j$, $j \in \{1, \cdots, |\mathbf{b}|\}$, such that $\tilde{c}_j = \tilde{c}_k$.
   (c) For every marked j:
      i. $\tilde{\mathbf{c}}_{\text{new}} \leftarrow \tilde{\mathbf{c}}'$;
      ii. $(\tilde{c}_{\text{new}})_j \leftarrow (\tilde{c}_{\text{new}})_j + \Delta c$;
      iii. $\mu_{\text{new}} \leftarrow \mu$;
      iv. $\mathbf{b}_{\text{new}} \leftarrow \mathbf{b}$;
      v. $\mu((\tilde{c}_{\text{new}})_j) \leftarrow \mu((\tilde{c}_{\text{new}})_j)+1$;
      vi. $\mathcal{S}_{\text{new}} \leftarrow \mathcal{S}_{\text{new}} \cup \{(\tilde{\mathbf{c}}_{\text{new}}, \mathbf{b}_{\text{new}}, \xi_{\text{new}}, \mu_{\text{new}})\}$.
4. If $|\hat{\mathbf{c}}(\mathbf{n})| < |\hat{\mathbf{c}}(\mathbf{n}')|$,
   (a) If $|\mathbf{b}| < m$,
      i. $\tilde{\mathbf{c}}_{\text{new}} \leftarrow (1 \quad \tilde{\mathbf{c}})$;
      ii. $\mathbf{b}_{\text{new}} \leftarrow (1 \quad \mathbf{b})$;
      iii. $\mu_{\text{new}} \leftarrow \begin{bmatrix} (1 \quad \tilde{\mathbf{c}}) = \tilde{\mathbf{c}}_{\text{new}} \\ \downarrow \quad \downarrow \mu \\ (1 \quad \mathbf{b}) = \mathbf{b}_{\text{new}} \end{bmatrix}$;
      iv. $\mathcal{S}_{\text{new}} \leftarrow \mathcal{S}_{\text{new}} \cup \{(\tilde{\mathbf{c}}_{\text{new}}, \mathbf{b}_{\text{new}}, \xi_{\text{new}}, \mu_{\text{new}})\}$.
   (b) If $|\mathbf{b}| > |\hat{\mathbf{c}}(\mathbf{n})|$, execute the same steps as 3a, 3b, and 3c.

5. If necessary, rearrange the elements of $\tilde{\mathbf{c}}_{\text{new}}$ (and update μ_{new} accordingly) such that $\tilde{\mathbf{c}}_{\text{new}}$ conforms to Definition 4.3.9.

6. Return $\mathcal{S}_{\text{new}}$.

4.4.3 Algorithm for computing an upper bound at arbitrary order

Now that we have introduced the major subroutines, we could put them together into an algorithm for finding a tight upper bound to $\|\mathbf{T}_r\|_\infty$, see Algorithm 4.

We start by recalling that Definition 4.3.10 can be thought of as the reduced configuration $\tilde{\mathbf{c}}$ space counterpart to the description of walks in the space of configuration $\mathbf{c}$ in Definition 4.3.7 in Section 4.3.4. We also define an energy combination $\mathbf{n}$ counterpart as the following.

Definition 4.4.2 (Walk in the space of energy combination n) *A sequence of energy combinations* $\mathbf{n}^{(0)} \to \mathbf{n}^{(1)} \to \cdots \to \mathbf{n}^{(r-1)} \to \mathbf{n}^{(r)}$ *is an* r-step walk in the space of energy combination $\mathbf{n}$ *(or* walk in $\mathbf{n}$ *for short) if*

$$\begin{aligned} &\mathbf{n}^{(0)} \in \mathcal{N}_-, \quad \mathbf{n}^{(r)} \in \mathcal{N}_- \\ &\mathbf{n}^{(i)} \in \mathcal{N}_+, \quad i = 1, \cdots, r-1. \end{aligned} \tag{4.53}$$

For every step from $\mathbf{n}^{(i)}$ *to* $\mathbf{n}^{(i+1)}$ *with* $i = 1, \cdots, r-2$*, either one of the following is true:*

1. $\mathbf{n}^{(i)} = \mathbf{n}^{(i+1)}$;

2. $\mathbf{n}^{(i+1)} = (n_0^{(i)}, n_1^{(i)}, \cdots, n_j^{(i)} - 1, n_{j+1}^{(i)} + 1, \cdots, n_{\ell-1}^{(i)})$.

For the initial step $\mathbf{n}^{(0)} \to \mathbf{n}^{(1)}$ *and final step* $\mathbf{n}^{(r-1)} \to \mathbf{n}^{(r)}$ *only case 2 above applies.*

The following definition concerns the step 2i and 2k of Algorithm 4, where the subroutine UpdateCell of Algorithm 3 is repeatedly invoked in all the cells of the automaton.

Algorithm 4: Algorithm for computing an upper bound to $\|\mathbf{T}_r\|_\infty$

Input: The order of perturbation, r, the scalar $\omega \in \mathbb{R}$ as in Definition 4.3.1, the vector $\boldsymbol{\lambda} \in \mathbb{R}^m$ as in Definition 4.3.2, the matrix $\mathbf{M} \in \mathbb{R}^{\ell \times \ell}$ as in Definition 4.3.3.

Output: An upper bound for $\|\mathbf{T}_r\|_\infty$, which we denote as τ_r.

Procedure $\tau_r = \textsc{PerturbBound}(r, \boldsymbol{\lambda}, \mathbf{M})$

1. Build the graph G_0 using Algorithm 2: $G_0(\mathcal{V}_0, \mathcal{E}_0) = \textsc{BuildCA}(|\boldsymbol{\lambda}|, \mathbf{M})$;
2. For each $\mathbf{n}_- \in \mathcal{N}_-$,
   (a) $G(\mathcal{V}, \mathcal{E}) \leftarrow G_0(\mathcal{V}_0, \mathcal{E}_0)$;
   (b) For all $\mathbf{n}$, $\mathbf{n}'$, $\mathcal{S}_{\mathbf{n}} \leftarrow \emptyset$ and $\mathcal{S}_{\mathbf{n},\mathbf{n}'} \leftarrow \emptyset$;
   (c) $\mathcal{E} \leftarrow \mathcal{E}_0 \backslash \bigcup_{\substack{\mathbf{n} \in \mathcal{N}_- \backslash \{\mathbf{n}_-\}, \\ \mathbf{n}' \in \mathcal{N}_-}} e(v_{\mathbf{n}}, v_{\mathbf{n}'})$; $\quad \mathcal{E}_{\text{dashed}} \leftarrow \bigcup_{\substack{\mathbf{n}' \in \mathcal{N}_+ \\ \mathbf{n} \in \mathcal{N}_-}} e(v_{\mathbf{n}'}, v_{\mathbf{n}})$;
   (d) $\mathcal{T}_- = \{\hat{\mathbf{c}}(\mathbf{n}_-), \quad \mathbf{b} = \underbrace{(0, \cdots, 0)}_{|\hat{\mathbf{c}}(\mathbf{n}_-)|}, \quad \xi = 1, \quad \mu : \hat{\mathbf{c}}(\mathbf{n}_-) \mapsto \mathbf{b}\}$;
   (e) Randomly choose a neighbor $\mathbf{n}_+ \in \mathcal{N}_+$ of $\mathbf{n}_-$;
   (f) Compute $\mathcal{S}_{\mathbf{n}_-,\mathbf{n}_+} = \textsc{Out}(\mathbf{n}_-, \mathbf{n}_+, \mathcal{T}_-)$ and randomly choose one 4-tuple $\mathcal{T}_{-+} \in \mathcal{S}_{\mathbf{n}_-,\mathbf{n}_+}$;
   (g) $\mathcal{S}_{\mathbf{n}_-,\mathbf{n}_+} \leftarrow \mathcal{T}_{-+}$;
   (h) $\mathcal{E} \leftarrow \mathcal{E} \backslash \mathcal{E}_{\text{dashed}}$;
   (i) Repeat $(r-1)$ times the following: For any $v_{\mathbf{n}} \in \mathcal{V}$, run $\textsc{UpdateCell}(v_{\mathbf{n}})$;
   (j) $\mathcal{E} \leftarrow \mathcal{E} \cup \mathcal{E}_{\text{dashed}}$;
   (k) For any $v_{\mathbf{n}} \in \mathcal{V}$, run $\textsc{UpdateCell}(v_{\mathbf{n}})$;
   (l) $\tau_{r,\mathbf{n}_-} \leftarrow \sum_{\mathbf{n} \in \mathcal{N}_-} \sum_{(\tilde{\mathbf{c}}, \mathbf{b}, \xi, \mu) \in \mathcal{S}_{\mathbf{n}}} \xi m_{\mathbf{b}}(\boldsymbol{\lambda})$;
3. Return $\tau_r = \max_{\mathbf{n}_- \in \mathcal{N}_-} \tau_{r,\mathbf{n}_-}$.

Definition 4.4.3 (Trace of the update algorithm) *Let $\mathcal{S}_{\mathbf{n}}^{(i)}$ and $\mathcal{S}_{\mathbf{n},\mathbf{n}'}^{(i)}$ be the set of 4-tuples associated with the node $v_{\mathbf{n}}$ and edge $e(v_{\mathbf{n}}, v_{\mathbf{n}'})$ respectively at the end of the i^{th} call to* UpdateCell *at step 2i of Algorithm 4. A* trace *of Algorithm 4 is a sequence of 4-tuples $\mathcal{T}^{(0)} \to \mathcal{T}^{(1)} \to \cdots \to \mathcal{T}^{(r)}$ that is associated with an r-step walk in the space of energy combinations $\mathbf{n}$ (or equivalently on the vertices of the graph $\mathbf{G}$ generated by* BuildCA *in Algorithm 2). The 4-tuple $\mathcal{T}^{(i)}$ at each step i is given by*

$$\mathcal{T}^{(i)} = \begin{cases} \mathcal{T}_{-+}, & i = 0 \\ \dfrac{1}{|z - E^{(i)}|}\textsc{Out}(\mathbf{n}^{(i-1)}, \mathbf{n}^{(i)}, \mathcal{T}^{(i-1)}), & i = 1, \cdots, r-1 \\ \textsc{Out}(\mathbf{n}^{(r-1)}, \mathbf{n}^{(r)}, \mathcal{T}^{(r-1)}), & i = r \end{cases} \tag{4.54}$$

where $\mathcal{T}_{-+}$ is computed by the initialization steps 2a through 2g of Algorithm 4.

From Equation 4.54 we see that $\mathcal{T}^{(i)} \in \mathcal{S}_{\mathbf{n}^{(i)}}^{(i)}$ for any $i = 0, \cdots, r$. Let $\mathcal{P}_r^{\tilde{\mathbf{c}}}$ be the set of r-step walks in the reduced configuration space (Definition 4.3.7) that starts from the initial reduced configuration $\tilde{\mathbf{c}}_1 = \emptyset$. Let $\mathcal{P}_r^{\mathcal{T}}$ be the set of r-step traces (Definition 4.4.3) generated by running UpdateCell procedure r times (Algorithm 3), with the initial input assigned by steps 2a through 2g of Algorithm 4. The following theorem shows that Algorithm 4 captures all the paths in the space of reduced configurations $\tilde{\mathbf{c}}$ that follow Definition 4.3.10.

Theorem 4.4.1 *There is a one-one correspondence (bijective mapping) between the two sets $\mathcal{P}_r^{\tilde{\mathbf{c}}}$ and $\mathcal{P}_r^{\mathcal{T}}$.*

Proof For every $k < r$, let $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$ be the set of k-step walks in the space of reduced configuration $\tilde{\mathbf{c}}$ obtained by truncating all r-step walks in $\mathcal{P}_r^{\tilde{\mathbf{c}}}$ at step k. There could be multiple walks in $\mathcal{P}_r^{\tilde{\mathbf{c}}}$ that share the same first k steps. We count them only once in $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$. Since $k < r$, every step of the k-step walks in $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$ is defined using Definition 4.4.3 but with all parts concerning $\mathbf{n}^{(r)}$ removed. Similarly, we define $\mathcal{Q}_k^{\mathcal{T}}$ as the set of k-step traces of the update algorithm obtained from truncating each trace in $\mathcal{P}_r^{\mathcal{T}}$ at the k^{th} step and counting the redundant elements only once.

To establish the theorem, we first show that for every $k < r$, there is a one-one correspondence between the elements of the two sets $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$ and $\mathcal{Q}_k^{\mathcal{T}}$. Specifically, for any k-step walk $q_k \in \mathcal{Q}_k^{\tilde{\mathbf{c}}}$ such that $q_k = \tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \to \tilde{\mathbf{c}}^{(k)}$, there is a trace of Algorithm 4 denoted as $t_k \in \mathcal{Q}_K^{\mathcal{T}}$, that can be described as $t_K = \mathcal{T}^{(0)} \to \mathcal{T}^{(1)} \to \cdots \to \mathcal{T}^{(k)}$ where $\mathcal{T}^{(i)} = (\tilde{\mathbf{c}}^{(i)}, \mathbf{b}^{(i)}, \xi^{(i)}, \mu^{(i)})$ for any $i \in \{1, \cdots, k\}$.

We use induction on k. For $k = 1$, $\mathcal{Q}_1^{\tilde{\mathbf{c}}} = \{\hat{\mathbf{c}}(\mathbf{n}_-)\}$ for some $\mathbf{n}_- \in \mathcal{N}_-$ (step 1 of Algorithm 4), which corresponds to $\mathcal{Q}_1^{\mathcal{T}} = \{\mathcal{T}_{-+}\}$. For the definition of $\mathcal{T}_{-+}$, refer to step 2f of Algorithm 4 respectively. By inspecting step 2a through 2g it is clear that the reduced energy configuration of $\mathcal{T}_{-+}$ is $\hat{\mathbf{c}}(\mathbf{n}_-)$. Hence the above statement is true for $k = 1$. Suppose the statement is true for all $k \leq K$. Then consider any K-step walk $q_K \in \mathcal{Q}_K^{\tilde{\mathbf{c}}}$ such that

$$q_K = \tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \to \tilde{\mathbf{c}}^{(K)}. \tag{4.55}$$

By induction hypodissertation, there must be a K-step trace $t_K \in \mathcal{Q}_K^{\mathcal{T}}$ that corresponds to p_K. Here the trace $t_K = \mathcal{T}^{(0)} \to \mathcal{T}^{(1)} \to \cdots \to \mathcal{T}^{(K)}$. It then suffices to show that all paths of the form $q'_{K+1} := q_K \to \tilde{\mathbf{c}}'$ has one-one correspondence with traces of the form $t_{K+1} = t_K \to \mathcal{T}_{\text{new}}$ where $\mathcal{T}_{\text{new}} = (\tilde{\mathbf{c}}', \mathbf{b}', \xi', \mu')$ is one of the new 4-tuples generated at either step 1 or 2 of UpdateCell in Algorithm 3. By Definition 4.3.10, $\tilde{\mathbf{c}}'$ has three possibilities:

(i) $\tilde{\mathbf{c}}' = \tilde{\mathbf{c}}^{(K)}$;

(ii) $|\tilde{\mathbf{c}}'| = |\tilde{\mathbf{c}}^{(K)}|$ and $|\tilde{\mathbf{c}}'_j - \tilde{\mathbf{c}}_j^{(K)}| = 1$ for some j;

(iii) $|\tilde{\mathbf{c}}'| = |\tilde{\mathbf{c}}^{(K)}| + 1$.

The case (i) is handled by the $\frac{\omega}{z - E(\mathbf{n})}\mathcal{S}_{\mathbf{n}}$ term in step 1 of Algorithm 3, with Equation 4.51. In other words, in this case q'_{K+1} maps to the trace $t_{K+1} := t_K \to \mathcal{T}_{\text{new}}$ with $\mathcal{T}_{\text{new}} = \frac{\omega}{z - E(\mathbf{n})}\mathcal{T}^{(K)} = (\tilde{\mathbf{c}}^{(K)}, \mathbf{b}^{(K)}, \frac{\omega}{z - E^{(K)}}\xi^{(K)}, \mu^{(K)})$ generated at step 1 of UpdateCell. Here $E^{(K)} = E(\mathbf{n}(\tilde{\mathbf{c}}^{(K)}))$.

The case (ii) is handled by steps 3 and 4b of Out in Algorithm 3. By definition, $\mathcal{T}^{(K)} = (\tilde{\mathbf{c}}^{(K)}, \mathbf{b}^{(K)}, \xi^{(K)}, \mu^{(K)})$. Recall $\tilde{\mathbf{c}}^{(K)} = (c_1^{(K)}, c_2^{(K)}, \cdots, c_{|\mathbf{b}^{(K)}|}^{(K)})$. Then $\tilde{\mathbf{c}}'$ is ob-

tained by incrementing or decrementing one of the ω_i elements by 1. Incrementing or decrementing any $\tilde{c}_i^{(K)}$ element will change the energy combination of $\tilde{\mathbf{c}}^{(K)}$. In particular, if there is a subset of the $\tilde{c}_i^{(K)}$ elements, call them $\tilde{c}_{i_1}^{(K)}$, $\tilde{c}_{i_2}^{(K)}$, $\cdots$, $\tilde{c}_{i_L}^{(K)}$, such that $\tilde{c}_{i_1}^{(K)} = \tilde{c}_{i_2}^{(K)} = \cdots = \tilde{c}_{i_L}^{(K)}$, then incrementing or decrementing any $\tilde{\mathbf{c}}_{i_j}^{(K)}$ could yield the same $\mathbf{n}(\tilde{\mathbf{c}}')$. In steps 3 and 4b we mark all such $\tilde{\mathbf{c}}_{i_j}^{(K)}$ elements. The case $\tilde{c}_{i_j}^{(K)} = 0$ for any $j = \{1, \cdots, L\}$ is handled in step 3 and the case $\tilde{c}_{i_j}^{(K)} \neq 0$ for any $j = \{1, \cdots, L\}$ is handled in step 4b. In either cases, the new 4-tuple $\mathcal{T}_{\text{new}} = (\tilde{\mathbf{c}}', \mathbf{b}', \xi', \mu')$ generated by OUT is such that we map the path $q'_{K+1} := q_K \to \tilde{\mathbf{c}}'$ to the trace $t_{K+1} := t_K \to \mathcal{T}_{\text{new}}$.

The case (iii) is handled in step 4a of OUT in Algorithm 3. In this case an inactive subsystem is active from E_0 to E_1. Hence $\tilde{\mathbf{c}}' = (1 \quad \tilde{\mathbf{c}}^{(K)})$. q'_{K+1} then maps to $t_{K+1} := t_K \to \mathcal{T}_{\text{new}}$ where $\mathcal{T}_{\text{new}} = (\tilde{\mathbf{c}}', \mathbf{b}', \xi', \mu')$ is generated at step 4a.

In summary we have shown that for each possible path $q'_{K+1} := q_K \to \tilde{\mathbf{c}}'$ in the reduced configuration space there is a corresponding trace of the algorithm $t_{K+1} := t_k \to \mathcal{T}_{\text{new}}$ where $\mathcal{T}_{\text{new}} = (\tilde{\mathbf{c}}', \mathbf{b}', \xi', \mu')$ is generated at various steps of Algorithm 3. Because these steps are at mutually exclusive branches of IF conditions, no q'_{K+1} maps to two different t_{K+1}'s simultaneously and vice versa. We also note that by Definition 4.3.10, q_K in Equation 4.55 must satisfy $\mathbf{n}(\tilde{\mathbf{c}}^{(i)}) \in \mathcal{N}_+$ for all $i = 1, \cdots, K$. This is enforced by step 2h in Algorithm 4, where all edges that goes from $\mathcal{N}_+$ to $\mathcal{N}_-$, namely the "dashed" edges, are removed.

We have thus far shown that for every walk in $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$, there is a corresponding trace in $\mathcal{Q}_k^{\mathcal{T}}$ that maps to it, and this is true for any $k < r$. Conversely, since any new triple $\mathcal{T}^{(i+1)}$ generated by $\mathcal{T}^{(i)}$ comes from either step 1 of UPDATECELL in Algorithm 3, or step 3 or 4a or 4b of OUT in Algorithm 3, and the cases ((i)), ((ii)) and ((iii)) above has accounted for each of the steps, we conclude that for every trace in $\mathcal{Q}_k^{\mathcal{T}}$ there must be a corresponding walk in $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$. Hence there is a one-one correspondence between the two sets $\mathcal{Q}_k^{\mathcal{T}}$ and $\mathcal{Q}_k^{\tilde{\mathbf{c}}}$ for any $k < r$.

By Definition 4.3.10 the final step of any r-step walk in the space of reduced configuration has to conform to the case 2 of Definition 4.3.10. Similarly, each trace in $\mathcal{P}_r^{\mathcal{T}}$ is associated with an r-step walk in the space of energy combination (Definition

4.4.2), for which the last step also needs to conform to the case 2 of Definition 4.4.2. Hence for any r-step walk in $\mathcal{P}_r^{\tilde{\mathbf{c}}}$, if the first $(r-1)$ steps are determined, the final step is also uniquely known. The same goes for any trace in $\mathcal{P}_r^{\mathcal{T}}$. We prove the theorem by using the one-one correspondence between $\mathcal{Q}_{r-1}^{\tilde{\mathbf{c}}}$ and $\mathcal{Q}_{r-1}^{\mathcal{T}}$ established from the previous inductive argument. The condition of returning to $\mathcal{N}_-$ at the last step, namely the restriction $\mathbf{n}(\tilde{\mathbf{c}}^{(r)}) \in \mathcal{N}_-$ in Equation 4.42 of Definition 4.3.10 is enforced in the step 2j of Algorithm 4 by adding back the dashed edges that enable transition from $\mathcal{N}_+$ back to $\mathcal{N}_-$. ■

From the above proof we could have a rough upper bound of the complexity of the algorithm. For a walk of r steps where each step has m choices, we have in total $O(m^r)$ possible walks. From the proof of Theorem 4.4.1 we have established that at any point during the algorithm, each 4-tuple at a node collects contributions from all possible walks up to the node. Hence at the r^{th} step of the algorithm, there are at most as many 4-tuples stored in all of the nodes as there are r-step walks. Each tuple takes $O(m)$ time to update since there are at most m elements of identical values in a reduced configuration $\tilde{\mathbf{c}}$ in case (ii) of OUT in Algorithm 3 while cases (i) and (iii) takes $O(1)$ time to treat. Putting these together, we have that r updates of the algorithm takes $O(rm^r)$. If we fix the order of perturbation r, this is polynomial with respect to the system size.

Theorem 4.4.1 shows that Algorithm 4 captures all the walks in $\tilde{\mathbf{c}}$ that conform to Definition 4.3.10. The theorem below shows that Algorithm 4 indeed computes the right hand side of Equation 4.43.

Theorem 4.4.2 (Correctness of Algorithm 4)
Given an r-step trace $\mathcal{T}^{(0)} \to \mathcal{T}^{(1)} \to \cdots \mathcal{T}^{(r)}$ as described in Definition 4.4.3 and (according to Theorem 4.4.1) its associated r-step walk $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)} \to \cdots \tilde{\mathbf{c}}^{(r)}$ in the space of reduced energy configurations $\tilde{\mathbf{c}}$, let $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(r)}$ be an

r-step walk in $\mathbf{c}$ *such that* $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}^{(i)}$. *Each step of the trace can be written as* $\mathcal{T}^{(i)} = (\tilde{\mathbf{c}}^{(i)}, \mathbf{b}^{(i)}, \xi^{(i)}, \mu^{(i)})$. *Then we have*

$$\xi^{(r)} m_{\mathbf{b}^{(r)}}(\boldsymbol{\lambda}) = \left(\prod_{i=1}^{r-1} \frac{1}{|z - E^{(i)}|} \right) \cdot m_{\mathbf{b}}(\boldsymbol{\lambda}) \cdot \left(\prod_{j:\exists i, j \in \mathcal{F}_i} M_{c_i^{(j-1)}, c_i^{(j)}} \right) \cdot \omega^k \tag{4.56}$$

where the symbols involved in the right hand side expression Equation 4.56 are the same as those defined in Equation 4.43 of Lemma 4.3.6.

Proof The proof of Lemma 4.3.3 is based on r-step walks that follow Definition 4.3.7. In fact from the arguments outlined by Equations 4.39, 4.40 and 4.41 we could see that any such r-step walk in the space of configuration $\mathbf{c}$ truncated at step q, $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(q)}$, contributes a multiplicative factor in one of the terms in the upper bound of $\|\mathbf{T}_r\|_2$ (refer to the right hand side of Equation 4.43) that can be written as

$$f(\mathbf{c}^{(0)}, \mathbf{c}^{(1)}) \cdot \frac{1}{|z - E^{(1)}|} \cdot f(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) \cdots f(\mathbf{c}^{(q-2)}, \mathbf{c}^{(q-1)}) \cdot \frac{1}{|z - E^{(q-1)}|} \cdot f(\mathbf{c}^{(q-1)}, \mathbf{c}^{(q)}). \tag{4.57}$$

The first step of the walk in $\tilde{\mathbf{c}}$, $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)}$, falls into either case 2 or 3 of Definition 4.3.10. In either case, PERTURBBOUND in Algorithm 4 will produce $\mathcal{T}_{-+}$ (step 2f) with partition $\mathbf{b}^{(1)} = (1)$ and coefficient $\xi^{(1)} = M_{\mathbf{n}_-, \mathbf{n}_+}$, which is correct because by Lemma 4.3.5, steps in $\mathbf{c}$ that are consistent with $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)}$ in the sense that $\tilde{\mathbf{c}}(\mathbf{c}^{(0)}) = \tilde{\mathbf{c}}^{(0)}$ and $\tilde{\mathbf{c}}(\mathbf{c}^{(1)}) = \tilde{\mathbf{c}}^{(1)}$ are but the same step $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)}$ with different permutations of the m subsystems (or elements of $\mathbf{c}$). In other words, the multiplicative factor associated with the step in reduced configuration $\tilde{\mathbf{c}}^{(0)} \to \tilde{\mathbf{c}}^{(1)}$ can be written as[3]

$$\sum_{\pi: \mathbb{N}^m \mapsto \mathbb{N}^m} f(\pi(\mathbf{c}^{(0)}), \pi(\mathbf{c}^{(1)})) = \sum_{\pi:[m] \mapsto [m]} \lambda_{\pi(j)} M_{st} = m_{(1)}(\boldsymbol{\lambda}) \cdot M_{st} \tag{4.58}$$

where we assume that during the step from $\mathbf{c}^{(0)}$ to $\mathbf{c}^{(1)}$, the j^{th} subsystem makes a transition from $\mathcal{P}_s$ to $\mathcal{P}_t$. From Equation 4.58 we see that the initial partition is indeed

[3]Here we abuse the notation π to mean a generic permutation over m elements. When π acts on an integer it returns another integer that results from the permutation. When π is applied on a vector of size m it permutes the m elements.

(1). Since in this one-step process only the j^{th} subsystem is acted on, $\mathcal{F}_j = \{1\}$ and $\mathcal{F}_i = \emptyset$ for any $i \neq j$ (for the definition of $\mathcal{F}_j$ see Lemma 4.3.6). The multiplicative factor $M_{\mathbf{n}_-,\mathbf{n}_+}$ is determined during a call to AddEdge in BuildCell of Algorithm 2. Since $c_j^{(0)} = s$ and $c_j^{(1)} = t$, $n(\tilde{\mathbf{c}}^{(1)})_s = n(\tilde{\mathbf{c}}^{(0)})_s - 1$ and $n(\tilde{\mathbf{c}}^{(1)})_t = n(\tilde{\mathbf{c}}^{(0)})_t + 1$. Hence a call to AddEdge$(\mathbf{n}(\tilde{\mathbf{c}}^{(0)}), \mathbf{n}(\tilde{\mathbf{c}}^{(1)}))$ adds weight M_{st} to the edge between the node for $\mathbf{n}(\tilde{\mathbf{c}}^{(0)})$ and that for $\mathbf{n}(\tilde{\mathbf{c}}^{(0)})$. Because in the context of PerturbBound in Algorithm 4, $\mathbf{n}(\tilde{\mathbf{c}}^{(0)}) = \mathbf{n}_-$ and $\mathbf{n}(\tilde{\mathbf{c}}^{(1)}) = \mathbf{n}_+$, $M_{\mathbf{n}_-,\mathbf{n}_+} = M_{st} = \xi^{(1)}$. We have thus far shown that Equation 4.58 holds for $r = 1$.

Next we will use induction to show that for any $1 < q < r - 1$,

$$\xi^{(q)} m_{\mathbf{b}^{(q)}}(\boldsymbol{\lambda}) = \left(\prod_{i=1}^{q-1} \frac{1}{|z - E^{(i)}|} \right) \cdot m_{\mathbf{b}^{(q)}}(\boldsymbol{\lambda}) \cdot \left(\prod_{j:\exists i, j \in \mathcal{F}_i^{(q)}} M_{c_i^{(j-1)}, c_i^{(j)}} \right) \cdot \omega^{k_q} \tag{4.59}$$

where $\mathcal{F}_i^{(q)} = \{j = 1, \cdots, q | c_i^{(j-1)} \neq c_i^{(j)}\}$ and $k_q = q - \sum_{i=1}^m |\mathcal{F}_i^{(q)}|$. Let $\mathbf{f}^{(q)}$ be such that $f_i = |\mathcal{F}_i^{(q)}|$, then $\mathbf{b}^{(q)}$ denotes $\mathbf{f}^{(q)}$ with its elements sorted in non-descending order to follow Definition 4.2.2 for reduced partitions. With the same rearrangement that leads to Equation 4.45 from Equation 4.44, one could see that the right hand side of Equation 4.59 is equal to Expression 4.57.

We start the induction by assuming that there is a $Q < r - 1$ such that Equation 4.59 holds for any $q \leq Q$. Now consider the Q^{th} call to UpdateCell (Algorithm 3) during the step 2i of PerturbBound in Algorithm 4 on the node associated with the energy combination $\mathbf{n}(\tilde{\mathbf{c}}^{(Q)})$. Depending on the step $\tilde{\mathbf{c}}^{(Q)} \to \tilde{\mathbf{c}}^{(Q+1)}$ there are 3 possible scenarios according to Definition 4.3.10:

(i) $\tilde{\mathbf{c}}^{(Q)} = \tilde{\mathbf{c}}^{(Q+1)}$. In this case $\mathcal{T}^{(Q+1)} = \frac{\omega}{|z - E^{(Q)}|} \mathcal{T}^{(Q)}$ from step 1 of UpdateCell in Algorithm 3. None of the sets $\mathcal{F}_i^{(Q)}$ are changed so $\mathcal{F}_i^{(Q)} = \mathcal{F}_i^{(Q+1)}$ for all $i = 1, \cdots, m$ and $\mathbf{b}^{(Q)} = \mathbf{b}^{(Q+1)}$. Therefore

$$\begin{aligned} \xi^{(Q+1)} m_{\mathbf{b}^{(Q+1)}}(\boldsymbol{\lambda}) &= \frac{\omega}{|z - E^{(Q)}|} \xi^{(Q)} m_{\mathbf{b}^{(Q)}}(\boldsymbol{\lambda}) \\ &= \frac{\omega}{|z - E^{(Q)}|} \cdot \prod_{i=1}^{Q-1} \frac{m_{\mathbf{b}^{(Q)}}(\boldsymbol{\lambda})}{|z - E^{(i)}|} \cdot \prod_{j:\exists i, j \in \mathcal{F}_i^{(Q)}} M_{c_i^{(j-1)}, c_i^{(j)}} \cdot \omega^{k_Q} \end{aligned}$$

$$= \prod_{i=1}^{(Q+1)-1} \frac{m_{\mathbf{b}^{(Q+1)}}(\boldsymbol{\lambda})}{|z - E^{(i)}|} \cdot \prod_{j:\exists i, j \in \mathcal{F}_i^{(Q+1)}} M_{c_i^{(j-1)}, c_i^{(j)}} \cdot \omega^{k_{Q+1}}. \quad (4.60)$$

where $k_{Q+1} = Q + 1 - \sum_{i=1}^{m} |\mathcal{F}_i^{(Q+1)}|$. On the second line we used the inductive hypodissertation Equation 4.59 for $q = Q$. By Equation 4.60 we have established that Equation 4.59 is also true $q = Q + 1$.

(ii) $\tilde{\mathbf{c}}^{(Q)}$ and $\tilde{\mathbf{c}}^{(Q+1)}$ differ by 1 at one element. Consider a walk in $\mathbf{c}$ with $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}^{(i)}$. Let h be such that $|\mathbf{c}_h^{(Q+1)} - \mathbf{c}_h^{(Q)}| = 1$. Note that here we are concerned with the walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \mathbf{c}^{(Q)}$ in $\mathbf{c}$ instead of $\tilde{\mathbf{c}}$, which by similar arguments that lead to Equation 4.45 from 4.44 in Lemma 4.3.6, contributes a factor

$$\begin{aligned} & f(\mathbf{c}^{(0)}, \mathbf{c}^{(1)}) \cdot \frac{1}{|z - E^{(1)}|} \cdot f(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) \cdots \\ & \qquad \cdots f(\mathbf{c}^{(Q-2)}, \mathbf{c}^{(Q-1)}) \cdot \frac{1}{|z - E^{(Q-1)}|} \cdot f(\mathbf{c}^{(Q-1)}, \mathbf{c}^{(Q)}) \\ & = \left(\prod_{i=1}^{m} \frac{1}{|z - E^{(i)}|} \right) \cdot \left(\prod_{i=1}^{m} \lambda_i^{|\mathcal{F}_i^{(Q)}|} \right) \cdot \left(\prod_{j:\exists i, j \in \mathcal{F}_i^{(Q)}} M_{c_i^{(j-1)}, c_i^{(j)}} \right) \cdot \omega^{k_Q}. \end{aligned} \quad (4.61)$$

Applying the inductive hypodissertation for $q = Q$, we have that the walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(Q)} \to \mathbf{c}^{(Q+1)}$ contributes an upper bound

$$\begin{aligned} & f(\mathbf{c}^{(0)}, \mathbf{c}^{(1)}) \cdot \frac{1}{|z - E^{(1)}|} \cdot f(\mathbf{c}^{(1)}, \mathbf{c}^{(2)}) \cdots \\ & \qquad \cdots f(\mathbf{c}^{(Q-1)}, \mathbf{c}^{(Q)}) \cdot \frac{1}{|z - E^{(Q)}|} \cdot \underbrace{f(\mathbf{c}^{(Q)}, \mathbf{c}^{(Q+1)})}_{= \lambda_h M_{c_h^{(Q)}, c_h^{(Q+1)}}} \\ & = \left(\prod_{i=1}^{m} \frac{1}{|z - E^{(i)}|} \right) \cdot \left(\prod_{i=1}^{m} \lambda_i^{|\mathcal{F}_i^{(Q)}|} \right) \cdot \lambda_h \cdot \\ & \qquad \cdot \left(\prod_{j:\exists i, j \in \mathcal{F}_i^{(Q)}} M_{c_i^{(j-1)}, c_i^{(j)}} \right) \cdot M_{c_h^{(Q)}, c_h^{(Q+1)}} \cdot \omega^{k_Q}. \end{aligned} \quad (4.62)$$

Here in Equation 4.62 λ_h will merge with the $\lambda_h^{\mathcal{F}_h^{(Q)}}$ term in the product, producing $\lambda_h^{|\mathcal{F}_h^{(Q)}|+1}$. Since by definition of $\mathcal{F}_i^{(Q)}$, $\mathcal{F}_h^{(Q+1)} = \mathcal{F}_h^{(Q)} \cup \{Q+1\}$, $|\mathcal{F}_h^{(Q+1)}| =$

$|\mathcal{F}_h^{(Q)}| + 1$. Because $\tilde{\mathbf{c}}^{(Q)} \neq \tilde{\mathbf{c}}^{(Q+1)}$ in this case, $k_Q = k_{Q+1}$. Finally, because $(Q+1) \in \mathcal{F}_h^{(Q+1)}$, $M_{\mathbf{c}_h^{(Q)}, c_h^{(Q+1)}}$ merges into the product $\prod_{j:\exists i, j \in \mathcal{F}_i^{(Q)}} M_{c_i^{(j-1)}, c_i^{(j)}}$ and Expression 4.62 becomes

$$\left(\prod_{i=1}^{(Q+1)-1} \frac{1}{|z - E^{(i)}|} \right) \cdot \left(\prod_{i=1}^{m} \lambda_i^{|\mathcal{F}_i^{(Q+1)}|} \right) \cdot \left(\prod_{j:\exists i, j \in \mathcal{F}_i^{(Q+1)}} M_{c_i^{(j-1)}, c_i^{(j)}} \right) \cdot \omega^{k_{Q+1}}. \quad (4.63)$$

By Lemma 4.3.5, the contribution of the $(Q+1)$-step walk in the reduced configuration $\tilde{\mathbf{c}}$ can be obtained by summing over all permutation $\pi : [m] \mapsto [m]$ of the subsystems, yielding

$$\left(\prod_{i=1}^{(Q+1)-1} \frac{1}{|z - E^{(i)}|} \right) \cdot \underbrace{\sum_{\pi:[m]\mapsto[m]} \left(\prod_{i=1}^{m} \lambda_{\pi(i)}^{|\mathcal{F}_{\pi(i)}^{(Q+1)}|} \right)}_{m_{\mathbf{b}^{(Q+1)}}(\boldsymbol{\lambda})} \cdot \left(\prod_{j:\exists i, j \in \mathcal{F}_i^{(Q+1)}} M_{c_i^{(j-1)}, c_i^{(j)}} \right) \cdot \omega^{k_{Q+1}}. \quad (4.64)$$

We would like to show that our Algorithm indeed computes expression 4.64 correctly. First of all, $\mathcal{T}^{(Q+1)}$ could only be generated by first making a function call to $\textsc{Out}(\mathbf{n}^{(Q)}, \mathbf{n}^{(Q+1)}, \mathcal{T}^{(Q)})$ in Algorithm 3. During the Out call, the algorithm starts out by reconstructing $\tilde{\mathbf{c}}^{(Q)}$ and $\tilde{\mathbf{c}}^{(Q+1)}$ from $\mathbf{b}^{(Q)}$, $\mathbf{n}(\tilde{\mathbf{c}}^{(Q)})$ and $\mathbf{n}(\tilde{\mathbf{c}}^{(Q+1)})$ at step 2 of Out. Since we assumed that $|c_h^{(Q+1)} - c_h^{(Q)}| = 1$, there must be an h' such that $|\tilde{c}_{h'}^{(Q+1)} - \tilde{c}_{h'}^{(Q)}| = 1$. In other words, $\tilde{c}_{h'}^{(Q+1)} = c_h^{(Q+1)}$, $\tilde{c}_{h'}^{(Q)} = c_h^{(Q)}$. The algorithm marks all such possible h' indices in $\tilde{\mathbf{c}}^{(Q)}$. With mapping $\mu^{(Q)}$ we are able to locate the element in $\mathbf{b}^{(Q)}$ that stores $|\mathcal{F}_h^{(Q)}|$. The algorithm Out then correctly increments the element by 1, to generate $\mathbf{b}^{(Q+1)}$. Because of the way AddEdge in Algorithm 2 is set up for constructing the cellular automaton which leads to $M_{\tilde{c}_{h'}^{(Q)}, \tilde{c}_{h'}^{(Q+1)}} = M_{\mathbf{n}^{(Q)}, \mathbf{n}^{(Q+1)}}$, the step 1 leads to $\xi^{(Q)} \leftarrow \xi^{(Q)} \cdot M_{\tilde{c}_{h'}^{(Q)}, \tilde{c}_{h'}^{(Q+1)}}$. Putting these together, we can see that $\textsc{Out}(\mathbf{n}^{(Q)}, \mathbf{n}^{(Q+1)}, \mathcal{T}^{(Q)})$ produces a 4-tuple

$$\mathcal{T}^{(Q,Q+1)} = (\tilde{\mathbf{c}}^{(Q+1)}, \mathbf{b}^{(Q+1)}, \xi^{(Q)} \cdot M_{c_h^{(Q)}, c_h^{(Q+1)}}, \mu^{(Q+1)}) \quad (4.65)$$

where $\mu^{(Q)} = \mu^{(Q+1)}$ because no new element is introduced in $\mathbf{b}^{(Q)}$. Then during step 1 of UpdateCell in Algorithm 3, $\mathcal{T}^{(Q+1)}$ is finally generated by the operation

$$\mathcal{T}^{(Q+1)} = \frac{1}{|z - E^{(Q)}|}\mathcal{T}^{(Q,Q+1)} = (\tilde{\mathbf{c}}^{(Q+1)}, \mathbf{b}^{(Q+1)}, \xi^{(Q+1)}, \mu^{(Q+1)}) \in \mathcal{S}_{\mathbf{n}(\tilde{\mathbf{c}}^{(Q)}),\mathbf{n}(\tilde{\mathbf{c}}^{(Q+1)})}. \tag{4.66}$$

It is straightforward to verify that $\xi^{(Q+1)} m_{\mathbf{b}^{(Q+1)}}(\boldsymbol{\lambda})$ equals to expression 4.64.

(iii) $|\tilde{\mathbf{c}}^{(Q+1)}| = |\tilde{\mathbf{c}}^{(Q)}| + 1$. Consider the same walk $\mathbf{c}^{(0)} \to \mathbf{c}^{(1)} \to \cdots \to \mathbf{c}^{(Q)} \to \mathbf{c}^{(Q+1)}$ as the case (ii) with $\tilde{\mathbf{c}}(\mathbf{c}^{(i)}) = \tilde{\mathbf{c}}^{(i)}$. In this case there is some h such that $\tilde{c}_h^{(Q)} = 0$ and $\tilde{c}_h^{(Q+1)} = 1$. Also $|\mathcal{F}_h^{(Q)}| = 0$ and $|\mathcal{F}_h^{(Q+1)}| = 1$. In other words a new element is added to $\mathbf{b}^{(Q)}$ to store $|\mathcal{F}_h^{(Q+1)}|$. Hence $\mathbf{b}^{(Q)}$ to store $|\mathcal{F}_h^{(Q+1)}|$. Hence $\mathbf{b}^{(Q+1)} = (\mathbf{b}^{(Q)} \quad 1)$. The algorithm identifies this case by testing if both $|\hat{\mathbf{c}}(\mathbf{n}(\mathbf{c}^{(Q)}))| < |\hat{\mathbf{c}}(\mathbf{n}(\tilde{\mathbf{c}}^{(Q+1)}))|$ and $|\mathbf{b}^{(Q)}| < m$ are true, because if the former is false it implies that $\tilde{\mathbf{c}}^{(Q+1)}$ has one more active element (Definition 4.3.8) with energy E_0 than $\tilde{\mathbf{c}}^{(Q)}$, which is impossible for any possible step $\tilde{\mathbf{c}}^{(Q)} \to \tilde{\mathbf{c}}^{(Q+1)}$ as stated in Definition 4.3.10. If $|\mathbf{b}^{(Q)}| = m$ then there is no h such that $|\mathcal{F}_h^{(Q)}| = 0$, another contradiction. Therefore the algorithm correctly captures the necessary and sufficient condition for this case and once it does, during the Q^{th} call to UpdateCell on $v_{\mathbf{n}(\tilde{\mathbf{c}}^{(Q)})}$ it generates a new partition $\mathbf{b}^{(Q+1)} = (\mathbf{b}^{(Q)} \quad 1)$, according to step 4(a)ii of Out in Algorithm 3, and the new element "1" is mapped from $\tilde{c}_h^{(Q+1)}$. The call $\textsc{Out}(\mathbf{n}(\tilde{\mathbf{c}}^{(Q)}), \mathbf{n}(\tilde{\mathbf{c}}^{(Q+1)}), \mathcal{T}^{(Q)})$ returns a new 4-tuple

$$\mathcal{T}^{(Q,Q+1)} = (\tilde{\mathbf{c}}^{(Q+1)}, \mathbf{b}^{(Q+1)}, \xi^{(Q)} \cdot M_{01}, \mu^{(Q+1)}) \in \mathcal{S}_{\mathbf{n}(\tilde{\mathbf{c}}^{(Q)}),\mathbf{n}(\tilde{\mathbf{c}}^{(Q+1)})} \tag{4.67}$$

The step 1 during the $(Q+1)^{\text{th}}$ call to UpdateCell on $v_{\mathbf{n}(\tilde{\mathbf{c}}^{(Q+1)})}$ generates

$$\mathcal{T}^{(Q+1)} = \frac{1}{|z - E^{(Q)}|}\mathcal{T}^{(Q,Q+1)} \tag{4.68}$$

with $\xi^{(Q+1)} m_{\mathbf{b}^{(Q+1)}}(\boldsymbol{\lambda})$ being equal to Expression 4.59 with $q = Q + 1$.

At the final step $\mathbf{c}^{(r-1)} \to \mathbf{c}^{(r)}$ only case (ii) holds. The same arguments carry over here. This concludes our proof. ■

4.4.4 Dealing with infinity

Obviously, computing the error exactly requires summing the perturbative series (Equation 4.15) to infinite order, which is not possible. Hence we make a relaxation by truncating the summation at some finite order p and proving that the norm of the sum from $p+1$ to infinity is bounded from above by some quantity that is easy to calculate. In particular, at p^{th} order, $p \geq 2$, we have the perturbative term $\mathbf{T}_p = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{p-2}\mathbf{G}_+\mathbf{V}_{+-}$. Suppose we have found an upper bound γ_p such that $\|\mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{p-2}\mathbf{G}_+\mathbf{V}_{+-}\| \leq \gamma_p$. Then an upper bound for the $p+1$-st order can be established using the inequality $\|\mathbf{AB}\| \leq \|\mathbf{A}\| \cdot \|\mathbf{B}\|$ for submultiplicative norms:

$$\|\mathbf{T}_{p+1}\| \leq \|\mathbf{T}_p\| \cdot \|\mathbf{G}_+\mathbf{V}_+\| \leq \gamma_p \|\mathbf{G}_+\mathbf{V}_+\| \leq \frac{\gamma_p}{\Delta}\|\mathbf{V}_+\|. \tag{4.69}$$

Here in the last inequality we have used the definition of Δ being the lowest excited state energy in the unperturbed Hamiltonian $\mathbf{H}$. Let $r = \|\mathbf{V}_+\|/\Delta$. Then we could bound the infinite sum by the triangle inequality:

$$\|\sum_{j=p+1}^{\infty} \mathbf{T}_j\| \leq \sum_{j=p+1}^{\infty} \|\mathbf{T}_j\| \leq \gamma_p(r + r^2 + r^3 + ...) = \gamma_p \frac{r}{1-r}. \tag{4.70}$$

To make sure that the series on the right hand side converges, we need $r < 1$, which is true for all the constructions we consider here.

4.5 Bit-flip gadgets: an example

In this section we use the gadget construction proposed in [7] and generalized in [9] to demonstrate the effectiveness of our algorithm. The gadget construction is called "bit-flip gadgets" in [1]. In Section 4.5.1 we show how our algorithm performs in analyzing the perturbative expansion associated with a particular 11-spin gadget Hamiltonian that simulates a target Hamiltonian consisting of two 3-local terms. We numerically show that the error estimation produced with our algorithm is essentially tight with respect to the exact value, which takes an exponential amount of computa-

tion to compute[4]. In Section 4.5.2 we show rigorously that with minor modification to the Algorithms in Section 4.4, we are able to efficiently compute perturbative terms at arbitrary order for the bit-flip gadgets.

4.5.1 An 11-spin gadget Hamiltonian

Consider the quantum system of 11 spins described in Figure 3a of the main text. The Hamiltonian can be expressed in form of the general setting $\widetilde{\mathbf{H}} = \mathbf{H}+\mathbf{V}$ described in Figure 1a of the main text. Here the unperturbed Hamiltonian $\mathbf{H}$ and perturbation $\mathbf{V}$ are defined as

$$\begin{aligned}
\mathbf{H} = \mathbf{H}^{(1)} + \mathbf{H}^{(2)}, \qquad & \mathbf{H}^{(1)} = \frac{\Delta}{4}(\mathbf{Z}_{u_1}\mathbf{Z}_{u_2} + \mathbf{Z}_{u_2}\mathbf{Z}_{u_3} + \mathbf{Z}_{u_1}\mathbf{Z}_{u_3}) \\
& \mathbf{H}^{(2)} = \frac{\Delta}{4}(\mathbf{Z}_{v_1}\mathbf{Z}_{v_2} + \mathbf{Z}_{v_2}\mathbf{Z}_{v_3} + \mathbf{Z}_{v_1}\mathbf{Z}_{v_3}) \\
\mathbf{V} = \mathbf{V}^{(1)} + \mathbf{V}^{(2)}, \qquad & \mathbf{V}^{(1)} = \mu_1(\mathbf{X}_1\mathbf{X}_{u_1} + \mathbf{X}_2\mathbf{X}_{u_2} + \mathbf{X}_3\mathbf{X}_{u_3}) \\
& \mathbf{V}^{(2)} = \mu_2(\mathbf{Y}_4\mathbf{X}_{v_1} + \mathbf{X}_2\mathbf{X}_{v_2} + \mathbf{Z}_5\mathbf{X}_{v_3})
\end{aligned} \tag{4.71}$$

where spins with u_i and v_i labels belong to the two unperturbed subsystems. Here we let Δ be orders of magnitude larger than μ_1 and μ_2 and keep the coefficients μ_1 and μ_2 as

$$\mu_1 = \left(\frac{\alpha_1\Delta^2}{6}\right)^{1/3}, \qquad \mu_2 = \left(\frac{\alpha_2\Delta^2}{6}\right)^{1/3} \tag{4.72}$$

where α_1 and α_2 are parameters related to the low energy effective Hamiltonian (see Equation 4.77). In Figure 3c of the main text we explicitly partition the Hamiltonian in the form of general setting discussed in Section 4.2.1 (Figure 1a of the main text).

The low-energy subspace of the total Hamiltonian $\widetilde{\mathbf{H}}$ is then $\mathcal{L}_- = \mathcal{L}_-^{(1)} \otimes \mathcal{L}_-^{(2)}$. Inspecting the expressions $\mathbf{H}^{(1)}$ and $\mathbf{H}^{(2)}$ gives the low energy subspaces for each subsystem: $\mathcal{L}_-^{(1)} = \text{span}\{|000\rangle_{u_1u_2u_3}, |111\rangle_{u_1u_2u_3}\}$ and $\mathcal{L}_-^{(2)} = \text{span}\{|000\rangle_{v_1v_2v_3}, |111\rangle_{v_1v_2v_3}\}$.

[4]In fact we choose 11 spins and no more because diagonalizing $2^{11} \times 2^{11}$ matrices is coming close to the RAM limit of the laptop computers being used.

For each subsystem $i \in \{1, 2\}$, the subspaces of $\mathbf{H}^{(i)}$ and their corresponding energies are

$$\begin{aligned} \mathcal{P}_0 &= \text{span}\{|000\rangle\}, & E_0 &= 0 \\ \mathcal{P}_1 &= \text{span}\{|001\rangle, |010\rangle, |100\rangle\}, & E_1 &= \Delta \\ \mathcal{P}_2 &= \text{span}\{|011\rangle, |101\rangle, |110\rangle\}, & E_2 &= \Delta \\ \mathcal{P}_3 &= \text{span}\{|111\rangle\}, & E_3 &= 0. \end{aligned} \tag{4.73}$$

In Figure 3d of the main text we show the spectrum of each subsystem. The vector $\boldsymbol{\lambda} = (\lambda_1, \lambda_2)$, which characterizes the "magnitudes" of perturbations onto each subsystem (Definition 4.3.2) can be determined based on Equation 4.71 as

$$\lambda_1 = \mu_1, \qquad \lambda_2 = \mu_2. \tag{4.74}$$

From the diagram in Figure 3d of the main text we could also determine the matrix $\mathbf{M}$ (see Definition 4.3.3) for this system. One could compute the matrix elements M_{ij} from the figure, where M_{ij} is the maximum, over all eigenstates of $\mathbf{H}$ in $\mathcal{P}_i$, number of possible transitions from a particular $|u\rangle \in \mathcal{P}_i$ to an eigenstate in $\mathcal{P}_j$. Precisely,

$$M_{ij} = \max_{|u\rangle \in \mathcal{P}_i} Card\{|v\rangle \in \mathcal{P}_j | \| \langle v|\mathbf{V}|u\rangle \| \neq 0\} \tag{4.75}$$

where $Card\{\cdot\}$ stands for cardinality (number of distinct elements) of a set. We could then determine that

$$\mathbf{M} = \begin{array}{c} \\ \mathcal{P}_0 \\ \mathcal{P}_1 \\ \mathcal{P}_2 \\ \mathcal{P}_3 \end{array} \begin{array}{c} \begin{array}{cccc} \mathcal{P}_0 & \mathcal{P}_1 & \mathcal{P}_2 & \mathcal{P}_3 \end{array} \\ \begin{bmatrix} & 3 & & \\ 1 & & 2 & \\ & 2 & & 1 \\ & & 3 & \end{bmatrix} \end{array} \tag{4.76}$$

where the row and column indices start from 0 because the subspaces $\mathcal{P}_0$, $\mathcal{P}_1$, $\cdots$, have indices that start from 0.

From Figure 3a and 3c of the main text we can see that the unperturbed system $\mathbf{H}$ essentially consists of two identical 4-level systems with energy levels E_0, E_1, E_2 and E_3. This gives rise to in total 9 possible energy combinations (Definition 4.3.5). Starting from the all-zero energy combination $\mathbf{n}_0 = [2, 0, 0, 0]$ and running Algorithm 2, we could construct a cellular automaton as shown in Figure 4d of the main text. We tabulate all the cells and their relevant information as in Figure 4c of the main text.

With the vector $\boldsymbol{\lambda}$ and the matrix $\mathbf{M}$ worked out as in Equations (4.74) and (4.76), we could use Algorithm 4 to find a tight upper bound for $\|\mathbf{T}_r\|_\infty$ at any order r. After a certain order p, when the upper bound becomes sufficiently small (assuming $\|\mathbf{T}_r\|_\infty \to 0$ as $r \to \infty$), we use Equation 4.70 to bound the terms from $p+1$ to infinity.

Using the perturbation series in Equation (4.15) we could show that if we truncate the series at the 3rd order, namely $\boldsymbol{\Sigma}_-(z) = \mathbf{H}_{\text{eff}} + \mathbf{T}_4 + \mathbf{T}_5 + \cdots$, we have the effective 3-body Hamiltonian

$$\mathbf{H}_{\text{eff}} = \alpha_1 \mathbf{X}_1 \mathbf{X}_2 \mathbf{X}_3 + \alpha_2 \mathbf{X}_2 \mathbf{Y}_4 \mathbf{Z}_5 + \gamma \mathbf{I} \tag{4.77}$$

for some γ that signifies the magnitude of the spectral shift. Here we let $\alpha_1 = 0.1$ and $\alpha = 0.2$. Then the entire Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ in Equation 4.71 is only dependent on a free parameter Δ. In order to test our algorithm for bounding perturbative terms, we treat terms from 4th order onward as errors in the perturbation series. This amounts to estimating $\|\boldsymbol{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|$. We could compute this value by explicitly computing $\boldsymbol{\Sigma}_-(z)$ by its definition $z\mathbf{I} - (\tilde{\mathbf{G}}_-(z))^{-1}$ and then evaluating $\|\boldsymbol{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|$. This method is inefficient but yields an accurate estimation for the error $\|\boldsymbol{\Sigma}_-(z) - \mathbf{H}_{\text{eff}}\|$. We will use it as a benchmark for comparison with the upper bound computed by the new algorithm developed here. As shown Figure 5 of the main text, the upper bounds computed by the cellular automaton algorithms are tight with respect to the exact calculation. For the purpose of comparison we also

compute the error bound due to triangle inequality (see Equation 4.18). We explicitly computed $\|\mathbf{V}\|_2$ and bounded $\|\mathbf{G}_+\|$ from above by $1/E_1$. Hence the simple bound based on Equation 4.18 becomes $\sum_{r=4}^{\infty} \|\mathbf{V}\|_2^r / E_1^{r-1} = \|\mathbf{V}\|_2^4 / (E_1^2 (E_1 - \|\mathbf{V}\|_2))$.

Note from Figure 5 of the main text that our upper bound based on the output of the CA algorithm only differs from the simple bound by a constant factor. This provides empirical justification for the method to treat infinity described in Section 4.4.4. When implementing the CA algorithm for the numerical example concerned in this section, we compute $\tau_r = \textsc{PerturbBound}(r, \boldsymbol{\lambda}, \mathbf{M})$ for r from 4 to a value p such that $\tau_p \leq 10^{-20}$. Then we resort to Equation 4.70 for computing an upper bound to $\|\mathbf{T}_{p+1} + \cdots\|_2$.

4.5.2 Rigorous arguments for the tightness of our error bound

Here we rigorously show that for perturbative gadget Hamiltonian proposed by Jordan and Farhi [9], we could efficiently compute terms $\mathbf{T}_r$ at arbitrary order. The purpose of the gadget construction is such that its low energy effective Hamiltonian is a sum of k-body interactions:

$$\mathbf{H}_{\text{eff}} = \sum_{i=1}^{m} c_i \mathbf{H}_{\text{eff},i} \tag{4.78}$$

where for now we assume that $c_i > 0$,

$$\mathbf{H}_{\text{eff},i} = \boldsymbol{\sigma}_{i,1} \boldsymbol{\sigma}_{i,2} \cdots \boldsymbol{\sigma}_{i,k}$$

and $\boldsymbol{\sigma}_{i,j} \in \{\mathbf{I}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ is either a Pauli or identity operator acting on the j^{th} qubit that the i^{th} term in $\mathbf{H}_{\text{eff}}$ acts on. An important assumption that we make here is that *for any* $i, j \in [m]$, $\mathbf{H}_{\textit{eff},i}$ *and* $\mathbf{H}_{\textit{eff},j}$ *commute*. In the notation introduced in Section 4.2.1, the gadget Hamiltonian $\widetilde{\mathbf{H}} = \mathbf{H} + \mathbf{V}$ takes the form

$$\begin{aligned} \mathbf{H} &= \sum_{i=1}^{m} \mathbf{H}^{(i)} \\ \mathbf{V} &= \sum_{i=1}^{m} \mathbf{V}^{(i)}. \end{aligned} \tag{4.79}$$

where

$$\begin{aligned} \mathbf{H}^{(i)} &= \sum_{1\le s<t\le k} \frac{1}{2}(\mathbf{I} - \mathbf{Z}_{i,s}\mathbf{Z}_{i,t}) \\ \mathbf{V}^{(i)} &= \sum_{j=1}^{k} |c_i|^{1/k} \boldsymbol{\sigma}_{i,j} \otimes \mathbf{X}_{i,j}. \end{aligned} \tag{4.80}$$

Here the coefficients in $\mathbf{V}^{(i)}$ are defined in a slightly different way from [9].

For each $\mathbf{H}_{\text{eff},i}$ there is a corresponding register of k extra spins (ancilla qubits) and $\mathbf{X}_{i,j}$, $\mathbf{Z}_{i,j}$ represent Pauli X and Z operators acting on the j^{th} ancilla accosiated with the i^{th} term in $\mathbf{H}_{\text{eff}}$. The spectrum of $\mathbf{H}^{(i)}$ is easy to find: the subspace of states with j qubits in $|1\rangle$ has energy $j(k-j)$. Hence $E_j^{(i)} = j(k-j)$. The ground state subspace of each subsystem $\mathbf{H}^{(i)}$ is $\mathcal{L}_-^{(i)} = \text{span}\{|0\rangle^{\otimes k}, |1\rangle^{\otimes k}\}$ and following Section 4.2.1 we let $\mathcal{L}_- = \mathcal{L}_-^{(1)} \otimes \cdots \otimes \mathcal{L}_-^{(m)}$. Following Lemma 4.3.1, the self energy expansion $\boldsymbol{\Sigma}_-(z) = \mathbf{H}_- + \mathbf{V}_- + \sum_{r=2}^{\infty} \mathbf{T}_r$ (Section 4.2.2) for $\widetilde{\mathbf{H}}$ then has the structure where at any order,

$$\mathbf{T}_r = \sum_{\substack{S\subseteq[m],\\ |S|\le\lfloor r/k\rfloor}} \mathbf{O}_{S,r} \otimes \boldsymbol{\Pi}_S \tag{4.81}$$

and

$$\mathbf{O}_{S,r} = \begin{cases} \alpha_r \mathbf{I} & S = \emptyset \\ \beta_{S,r} \prod_{i\in S} \mathbf{H}_{\text{eff},i} & S \neq \emptyset \end{cases} \tag{4.82}$$

with α_r and $\beta_{S,r}$ being real coefficients, and

$$\boldsymbol{\Pi}_S = \prod_{i\in S} \left(\bigotimes_{j=1}^{k} |0\rangle\langle 1|_{i,j} + \bigotimes_{j=1}^{k} |1\rangle\langle 0|_{i,j} \right). \tag{4.83}$$

Based on the intuition about the connection between $\mathbf{T}_r$ and walks among the eigenstates of $\mathbf{H}$ driven by $\mathbf{V}$ (Definition 4.3.6), we could see that for any r, $\mathbf{T}_r$ could contain terms that are associated with two possible types of walks:

- Walks that start from $|0\rangle^{\otimes k}$ (resp. $|1\rangle^{\otimes k}$) and end at $|0\rangle^{\otimes k}$ (resp. $|1\rangle^{\otimes k}$);
- Walks that start from $|0\rangle^{\otimes k}$ (resp. $|1\rangle^{\otimes k}$) and end at $|1\rangle^{\otimes k}$ (resp. $|0\rangle^{\otimes k}$).

For $r < k$, only the first type of walk is possible since $\mathbf{V}$ flips one ancilla at a time and in this case the order of perturbation is simply not high enough for $|0\rangle^{\otimes k}$ to reach $|1\rangle^{\otimes k}$. Hence $S = \emptyset$ in this case and in Equation 4.82 the corresponding operator is proportional to identity. In case the order of perturbation is high enough to drive transitions between degenerate states in $\mathcal{L}_-$, the set S describes the subset of m subsystems for which a transition from $|0\rangle^{\otimes k}$ to $|1\rangle^{\otimes k}$ has occured. Hence the expression in Equation 4.82 for the case $S \neq \emptyset$. Note that this expression only applies when the $\mathbf{H}_{\text{eff},i}$ terms commute.

In accordance with Equation 4.82 we define a scalar function

$$f_r = |\alpha_r| + \sum_{\substack{S \subseteq [m] \\ 1 \leq |S| \leq \lfloor r/k \rfloor}} |\beta_{S,r}| = |\alpha_r| + \sum_{i=1}^{\lfloor r/k \rfloor} \gamma_{i,r} \tag{4.84}$$

where we define $\gamma_{i,r} = \sum_{S \subseteq [m], |S|=i} |\beta_{S,r}|$. Our goal is to show that

- Under the assumptions of pairwise commutativity between terms in $\mathbf{H}_{\text{eff}}$, we have $\|\mathbf{T}_r\|_2 = f_r$.
- Using the cellular automaton algorithms that we introduced, f_r can be computed in time $O(rm^r)$.

The combination of both goals will show that our cellular automaton algorithms essentially yield *tight* upper bounds for $\|\mathbf{T}_r\|_2$ if terms in $\mathbf{H}_{\text{eff}}$ commute. If the terms in $\mathbf{H}_{\text{eff}}$ do not commute, $\|\mathbf{T}_r\|_\infty < f_r$ strictly and the difference $|f_r - \|\mathbf{T}_r\|_\infty|$ depends on more detailed structure of non-commutativity among the terms of $\mathbf{H}_{\text{eff}}$.

To prove rigorously that $\|\mathbf{T}_r\|_2 = f_r$ when the $\mathbf{H}_{\text{eff},i}$ terms commute pairwise, we first note the following Lemma.

Lemma 4.5.1 *For any operators* $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{2^n \otimes 2^n}$ *such that*

$$\begin{aligned} \mathbf{A} &= \mathbf{P}_1 \otimes \mathbf{P}_2 \otimes \cdots \otimes \mathbf{P}_n \\ \mathbf{B} &= \mathbf{Q}_1 \otimes \mathbf{Q}_2 \otimes \cdots \otimes \mathbf{Q}_n \end{aligned} \tag{4.85}$$

where for any i, $\mathbf{P}_i, \mathbf{Q}_i \in \{\mathbf{I}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$. *Then for commuting* $\mathbf{A}$ *and* $\mathbf{B}$,

$$\|c_1\mathbf{A} + c_2\mathbf{B}\|_\infty = \begin{cases} |c_1 + c_2| & \text{if } \mathbf{A} = \mathbf{B} \\ |c_1| + |c_2| & \text{otherwise.} \end{cases} \tag{4.86}$$

Proof The case where $\mathbf{A} = \mathbf{B}$ is trivial so we focus on the case where $\mathbf{A} \neq \mathbf{B}$. Clearly $\|c_1\mathbf{A} + c_2\mathbf{B}\|_\infty \leq |c_1| + |c_2|$. On the other hand $\|c_1\mathbf{A} + c_2\mathbf{B}\|_\infty \geq \|c_1\mathbf{A} + c_2\mathbf{B}\|_2$. We show that $\|c_1\mathbf{A}+c_2\mathbf{B}\|_2 = |c_1|+|c_2|$ by first noting that $\mathbf{A}$ and $\mathbf{B}$ must differ in at least two operators. That is, there must be $i \neq j \in [m]$ such that $\mathbf{P}_i \neq \mathbf{Q}_i$ and $\mathbf{P}_j \neq \mathbf{Q}_j$. Then we choose a state $|\phi\rangle = |\phi_1\rangle \otimes |\phi_2\rangle \otimes \cdots \otimes |\phi_m\rangle$ where each $|\phi_s\rangle \in \mathbb{C}^2$ is a single-qubit state, such that $\langle\phi|c_1\mathbf{A} + c_2\mathbf{B}|\phi\rangle = |c_1| + |c_2|$. For example if $c_1 > 0$ and $c_2 < 0$, we will choose $|\phi_i\rangle$ such that $\langle\phi_i|\mathbf{P}_i|\phi_i\rangle = 1$ but $\langle\phi_i|\mathbf{Q}_i|\phi_i\rangle = -1$, which is possible considering $\mathbf{P}_i, \mathbf{Q}_i \in \{\mathbf{I}, \mathbf{X}, \mathbf{Y}, \mathbf{Z}\}$ and $\mathbf{P}_i \neq \mathbf{Q}_i$. We then choose for every $s \in [m]\backslash\{i\}$ we have $\langle\phi_s|\mathbf{P}_s|\phi_s\rangle = 1$ and $\langle\phi_s|\mathbf{Q}_s|\phi_s\rangle = 1$. This way $\langle\phi|c_1\mathbf{A} + c_2\mathbf{B}|\phi\rangle = |c_1| + |c_2|$. Similar idea works for other combinations of signs of c_1 and c_2. ■

Since $\|\mathbf{A}\|_\infty = \|\mathbf{A}\|_2 = 1$ and the same holds for $\mathbf{B}$ (Equation 4.85), Equation 4.86 in Equation 4.86 also applies to the 2-norm of $c_1\mathbf{A} + c_2\mathbf{B}$. For the projectors $\mathbf{\Pi}_S$ defined in Equation 4.83, we have

$$\|\mathbf{\Pi}_S\|_2 = \|\mathbf{\Pi}_S\|_\infty = \left\|\prod_{i\in S} \mathbf{X}_i\right\|_\infty = \left\|\prod_{i\in S} \mathbf{X}_i\right\|_2. \tag{4.87}$$

Combining Equation 4.87 with the expression of $\mathbf{O}_{S,r}$ in Equation 4.82 we see that the terms in the summation of Equation 4.81 are pairwise commutative. Applying Equation 4.86 and 4.87 yields the desired tightness result

$$\|\mathbf{T}_r\|_2 = \left\| \sum_{\substack{S\subseteq[m] \\ |S|\leq\lfloor r/k\rfloor}} \mathbf{O}_{S,r} \otimes \prod_{i\in S} \mathbf{X}_i \right\|_2 = |\alpha_r| + \sum_{\substack{S\subseteq[m] \\ 1\leq|S|\leq\lfloor r/k\rfloor}} |\beta_{S,r}| = f_r. \tag{4.88}$$

In general when the terms $\mathbf{H}_{\text{eff},i}$ do not commute, of course we always have the inequality $\|c_1\mathbf{A} + c_2\mathbf{B}\|_2 \leq \|c_1\mathbf{A} + c_2\mathbf{B}\|_\infty \leq |c_1| + |c_2|$. Hence $\|\mathbf{T}_r\|_2 \leq \|\mathbf{T}_r\|_\infty \leq f_r$, showing that f_r is always an upper bound to $\|\mathbf{T}_r\|_2$.

The norm $\|\mathbf{T}_r\|_\infty$ can be computed by explicit calculation of $\mathbf{T}_r$, which takes an *exponential* amount of computation as the system size m increases. Here we show that using a slightly modified variant of our cellular automaton algorithm we can compute f_r in time $O(rm^r)$, which is *polynomial* time in m. The exponential dependence on r in the runtime is related to the number of possible r-step Motzkin walks and is likely difficult to improve.

In order to compute f_r, we start by identifying the parameters ω, $\boldsymbol{\lambda}$ and $\mathbf{M}$ (Section 4.3.1) for the gadget Hamiltonian $\widetilde{\mathbf{H}}$. Because $\mathbf{V}$ always induces transitions on the eigenstates of $\mathbf{H}$, by Definition 4.3.1, $\omega = 0$. We define the elements of $\boldsymbol{\lambda} \in \mathbb{R}^m$ to be $\lambda_s = |c_s|^{1/k}$ for $s = 1, \cdots, m$. From the definition of $\mathbf{H}^{(s)}$ we could see that the elements of $\mathbf{M}$ should be assigned such that $M_{st} = k - s$ if $t = s + 1$, $M_{st} = s$ if $t = s - 1$, and 0 otherwise.

In addition, we make the following modifications to our algorithms:

1. In the PerturbBound subroutine of Algorithm 4, at step instead of looping over $\mathcal{N}_-$, fix $\mathbf{n}_- = \mathbf{n}_0 = (m, 0, \cdots, 0)$.

2. At step 2l, compute

$$\begin{aligned} \alpha_r &= \sum_{(\tilde{\mathbf{c}},\mathbf{b},\xi,\mu)\in\mathcal{S}_{\mathbf{n}_0}} \xi m_{\mathbf{b}}(\boldsymbol{\lambda}), \\ \gamma_{i,r} &= \sum_{(\tilde{\mathbf{c}},\mathbf{b},\xi,\mu)\in\mathcal{S}_{\mathbf{n}_i}} \xi m_{\mathbf{b}}(\boldsymbol{\lambda}), \quad i = 1, \cdots, \lfloor r/k \rfloor \end{aligned} \tag{4.89}$$

   where $\mathbf{n}_i = (m - i, 0, \cdots, 0, i)$.

3. At step 3, instead of τ_r, return f_r according to Equation 4.84.

The structure of $\mathbf{V}^{(s)}$ does satisfy the block-tridiagonal property assumed in Section 4.2.1. Therefore all the arguments on the walks in the space of $\mathbf{c}$, $\tilde{\mathbf{c}}$ and $\mathbf{n}$ hold. The parameters of f_r can be computed by calling $f_r = \text{PerturbBound}(r, \boldsymbol{\lambda}, \mathbf{M})$ with modifications listed above. The efficiency of this procedure is argued after the proof of Theorem 4.4.1.

4.6 Discussion and conclusion

- Our algorithms are constructed based on a physical setting that is not without assumptions. The first major assumption concerns the structure of $\mathbf{V}$ as described in Equation 4.11 and 4.12. The block tridiagonal structure of $\mathbf{V}^{(i)}$ has a direct consequence on what transitions are possible during one step of a walk, be it in $\mathbf{H}$ eigenstates (Definition 4.3.6), energy configuration $\mathbf{c}$ (Definition 4.3.7), reduced energy configuration $\tilde{\mathbf{c}}$ (Definition 4.3.10) or energy combination $\mathbf{n}$ (Definition 4.4.2). In case one would like to relax the assumption of $\mathbf{V}^{(i)}$ being tridiagonal and would like to instead treat $\mathbf{V}^{(i)}$'s that are band diagonal with the band width being greater than 3, the definitions of the walks will need to be modified to account for $\mathbf{V}$ being able to change an element of $\mathbf{c}$ by more than 1 during a single step $\mathbf{c}^{(i)} \to \mathbf{c}^{(i+1)}$. The algorithms will also need to be adjusted accordingly.

  A second assumption concerns the magnitude of $\mathbf{V}$. Here in order to guarantee the convergence of perturbation series $\mathbf{\Sigma}_{-}(z)$ in the regime of z specified by Theorem 4.2.1, we assume that $\|\mathbf{V}\|_2 \leq \Delta/2$. In general this assumption could be weakened [2] to a statement that ultimately is not dependent on any global property of $\mathbf{V}$, such as $\|\mathbf{V}\|_2$, and the series in $\mathbf{\Sigma}_{-}(z)$ still converges and Theorem 4.2.1 could still hold.

- We derive the upper bound using symmetric polynomials, as one could see from Lemma 4.3.3 and Lemma 4.3.6. An implicit assumption on using symmetric polynomials is that the terms in $\mathbf{V}$ commute with each other. Otherwise for example if $\mathbf{V}$ contains terms that are proportional to $\lambda_1 \mathbf{X}_i$ and $\lambda_2 \mathbf{Z}_i$ operating on the same spin i, at high orders one may expect terms such as $\lambda_1\lambda_2 \mathbf{X}_i\mathbf{Z}_i + \lambda_2\lambda_1 \mathbf{Z}_i\mathbf{X}_i$, which is vanishing but the symmetric polynomial would include such terms as $\lambda_1\lambda_2 + \lambda_2\lambda_1 = 2\lambda_1\lambda_2$, which is non-zero. This unawareness of non-commutativity will cause the upper bound computed by the algorithm to be less tight than the case shown in Section 4.5, where all terms in $\mathbf{V}$ commute.

- Perhaps one of the areas where our algorithm could find direct application is adiabatic quantum computation, where one often works with quantum systems with simple, restricted forms of interaction but wishes to realize some effective interactions $\mathbf{H}_{\text{eff}}$ that are more complicated. A common idea is to construct a Hamiltonian $\widetilde{\mathbf{H}}$ for which perturbation theory gives rise to $\mathbf{H}_{\text{eff}}$ at the first few orders. Then it becomes instrumental to have accurate estimation of how large the higher order error terms are. In fact a seemingly minor improvement in error estimation could lead to significant reduction in the resource required for producing $\mathbf{H}_{\text{eff}}$ using constructions of $\widetilde{\mathbf{H}}$, see for example [1]. Our algorithm certainly will enable improvement on a broader class of constructions of $\widetilde{\mathbf{H}}$ for adiabatic quantum computing than prior works by providing accurate error estimates that are not available with simple techniques (such as those that lead to Equation 4.18).

- The parallel nature of the update rules in cellular automata could facilitate parallelism in the software implementation of our algorithms, which will further speed up the computation. For example, with $O(m)$ processors each storing the information of one cell and its out going edges, the algorithm takes $O(rh(r))$ time. Here $h(r)$ is the maximum number of 4-tuples stored in any cell or edge during the algorithm.

LIST OF REFERENCES

LIST OF REFERENCES

[1] Y. Cao, R. Babbush, J. Biamonte, and S. Kais, "Hamiltonian gadgets with reduced resource requirements," *Physical Review A*, vol. 91, no. 1, p. 012315, 2015. arXiv:1311.2555 [quant-ph].

[2] Y. Cao and D. Nagaj, "Perturbative gadget without strong interactions," *Quantum Information and Computation*, vol. 15, pp. 1197–1222, 2015. arXiv:1408.5881 [quant-ph].

[3] D. Aharonov, A. Harrow, Z. Landau, D. Nagaj, M. Szegedy, and U. Vazirani, "Local tests of global entanglement and a counterexample to the generalized area law," in *IEEE 55th Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 246–255, 2014.

[4] Y. Cao and S. Kais, "Efficient estimation of perturbative error with cellular automata," 2016. arXiv:1607.01374 [quant-ph].

[5] D. Deutsch, "Quantum theory, the Church-Turing principle and the universal quantum computer," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 400, no. 1818, pp. 97–117, 1985.

[6] E. Bernstein and U. Vazirani, "Quantum complexity theory," *Society for Industrial and Applied Mathematics Journal of Computing*, vol. 26, pp. 1411–1473, Oct. 1997.

[7] J. Kempe, A. Kitaev, and O. Regev, "The complexity of the Local Hamiltonian problem," *Society for Industrial and Applied Mathematics Journal of Computing*, vol. 35, no. 5, pp. 1070–1097, 2006. arXiv:quant-ph/0406180.

[8] R. Oliveira and B. Terhal, "The complexity of quantum spin systems on a two-dimensional square lattice," *Quantum Information and Computation*, vol. 8, no. 10, pp. 0900–0924, 2008. arXiv:quant-ph/0504050.

[9] S. P. Jordan and E. Farhi, "Perturbative gadgets at arbitrary orders," *Physical Review A*, vol. 062329, Feb 2008.

[10] R. P. Feynman, "Simulating physics with computers," *International Journal of Theoretical Physics*, vol. 21, no. 6, pp. 467–488, 1982.

[11] R. P. Feynman, "Quantum mechanical computers," *Optics News*, vol. 11, pp. 11–20, Feb 1985.

[12] Y. Manin, *Computable and Uncomputable*. Moscow: Sovetskoye Radio, 1980.

[13] P. W. Shor, "Algorithms for quantum computation: Discrete logarithms and factoring," in *Proceedings of the 35th Annual Symposium on Foundations of Computer Science*, pp. 124–134, IEEE Computer Society, 1994.

[14] P. W. Shor, "Polynomial-time algorithms for prime factorization and discrete logarithms on a quantum computer," *Society for Industrial and Applied Mathematics Journal of Computing*, vol. 26, pp. 1484–1509, Oct. 1997.

[15] D. Aharonov and T. Naveh, "Quantum NP – A Survey," 2002. arXiv:quant-ph/0210077.

[16] J. Watrous, *Quantum Computational Complexity*, pp. 7174–7201. New York, NY, USA: Springer New York, 2009.

[17] S. Gharibian, Y. Huang, Z. Landau, and S. W. Shin, "Quantum Hamiltonian complexity," *Foundations and Trends in Theoretical Computer Science*, vol. 10, no. 3, pp. 159–282, 2015.

[18] Y.-K. Liu, "Consistency of Local Density Matrices is QMA-complete," in *Proceedings of the Nineth International Conference on Approximation Algorithms for Combinatorial Optimization Problems, and 10th International Conference on Randomization and Computation*, APPROX'06/RANDOM'06, pp. 438–449, Springer-Verlag, 2006.

[19] S. Aaronson, "Computational complexity: Why quantum chemistry is hard," *Nature Physics*, vol. 5, pp. 707–708, 2009.

[20] N. Schuch and F. Verstraete, "Computational complexity of interacting electrons and fundamental limitations of density functional theory," *Nature Physics*, vol. 5, 2009.

[21] J. D. Whitfield, M.-H. Yung, D. G. Tempel, S. Boixo, and A. Aspuru-Guzik, "Computational complexity of time-dependent density functional theory," *New Journal of Physics*, vol. 16, no. 083035, 2014.

[22] J. D. Whitfield, N. Schuch, and F. Verstraete, *The Computational Complexity of Density Functional Theory*, pp. 245–260. Springer International Publishing, 2014.

[23] F. Barahona, "On the computational complexity of Ising spin glass models," *Journal of Physics A: Mathematical and General*, vol. 15, no. 10, p. 3241, 1982.

[24] T. Albash, W. Vinci, A. Mishra, P. A. Warburton, and D. A. Lidar, "Consistency tests of classical and quantum models for a quantum annealer," *Physical Review A*, vol. 91, p. 042314, 2015.

[25] H. G. Katzgraber, F. Hamze, Z. Zhu, A. J. Ochoa, and H. Munoz-Bauza, "Seeking quantum speedup through spin glasses: The good, the bad, and the ugly," *Physical Review X*, vol. 5, p. 031026, 2015.

[26] I. Hen, J. Job, T. Albash, T. F. Ronnøw, M. Troyer, and D. Lidar, "Probing for quantum speedup in spin glass problems with planted solutions," *Physical Review A*, vol. 92, p. 042325, 2015.

[27] D. Venturelli, S. Mandrà, S. Knysh, B. O'Gorman, R. Biswas, and V. Smelyanskiy, "Quantum optimization of fully-connected spin glasses," *Physical Review X*, vol. 5, p. 031040, 2015.

[28] S. A. Cook, "The complexity of theorem-proving procedures," in *Proceedings of the Third Annual ACM Symposium on Theory of Computing*, pp. 151–158, ACM, 1971.

[29] L. Levin, "Universal search problems," *Problems of Information Transmission*, vol. 9, no. 3, pp. 115–116, 1973.

[30] B. A. Trakhtenbrot, "A survey of Russian approaches to perebor (brute-force searches) algorithms," *Annals of the History of Computing*, vol. 6, no. 4, 1984.

[31] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi, *Classical and quantum computation*. Providence, Rhode Island: American Mathematical Society, 2002.

[32] T. Cubitt and A. Montanaro, "Complexity classification of Local Hamiltonian problems," 2013. arXiv:1311.3161 [quant-ph].

[33] J. D. Biamonte and P. J. Love, "Realizable Hamiltonians for universal adiabatic quantum computers," *Physical Review A*, vol. 78, p. 012352, Jul 2008.

[34] A. Lascoux, *Symmetric Functions and Combinatorial Operators on Polynomials*. Providence, Rhode Island: American Mathematical Society, 2003.

[35] R. Shankar, *Principles of Quantum Mechanics*. New York, NY, USA: Plenum Press, 1994.

[36] D. J. Griffiths, *Introduction to Quantum Mechanics*. Upper Saddle River, NJ, USA: Pearson Prentice Hall, 2004.

[37] A. Messiah, *Quantum Mechanics*. New York, USA: Dover Publications Inc., 2014.

[38] S. Aaronson, *Quantum computing since Democritus*. Cambridge, UK: Cambridge University Press, 2013.

[39] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory*. Hoboken, NJ, USA.: John Wiley and Sons, 2000.

[40] T. Kato, *Perturbation theory for linear operators*. Springer-Verlag, 1966.

[41] M. Nielsen and I. Chuang, *Quantum computation and quantum information*. Cambridge, UK: Cambridge University Press, 2000.

[42] N. D. Mermin, *Quantum Computer Science*. Cornell University, New York: Cambridge University Press, 2007.

[43] E. Rieffel and W. Polak, *Quantum Computing: A Gentle Introduction*. Cambridge, MA, USA: MIT Press, 2011.

[44] D. Aharonov, W. van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev, "Adiabatic quantum computation is equivalent to standard quantum computation," *Society for Industrial and Applied Mathematics Journal of Computing*, vol. 37, pp. 166–194, 2007.

[45] C. Monroe, D. M. Meekhof, B. E. King, W. M. Itano, and D. J. Wineland, "Demonstration of a fundamental quantum logic gate," *Physical Review Letter*, vol. 75, pp. 4714–4717, Dec 1995.

[46] T. A. Brun and R. Schack, "Realizing the quantum baker's map on a NMR quantum computer," *Physical Review A*, vol. 59, pp. 2649–2658, Apr 1999.

[47] R. Marx, A. F. Fahmy, J. M. Myers, W. Bermel, and S. J. Glaser, "Realization of a 5-bit NMR quantum computer using a new molecular architecture," 1999. arXiv:quant-ph/9905087.

[48] L. M. K. Vandersypen, M. Steffen, G. Breyta, C. S. Yannoni, M. H. Sherwood, and I. L. Chuang, "Experimental realization of Shor's quantum factoring algorithm using nuclear magnetic resonance," *Nature*, vol. 414, pp. 883–887, 2001.

[49] C. Negrevergne, T. S. Mahesh, C. A. Ryan, M. Ditty, F. Cyr-Racine, W. Power, N. Boulant, T. Havel, D. G. Cory, and R. Laflamme, "Benchmarking quantum control methods on a 12-qubit system," *Physical Review Letter*, vol. 96, p. 170501, May 2006.

[50] T. Monz, P. Schindler, J. T. Barreiro, M. Chwalla, D. Nigg, W. A. Coish, M. Harlander, W. Hänsel, M. Hennrich, and R. Blatt, "14-qubit entanglement: Creation and coherence," *Physical Review Letter*, vol. 106, p. 130506, Mar 2011.

[51] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser, "Quantum computation by adiabatic evolution," *MIT-CTP-2936*, 2000.

[52] F. Barahona, "On the computational complexity of Ising spin glass models," *Journal of Physics A: Mathematical and General*, vol. 15, no. 10, p. 3241, 1982.

[53] A. Lucas, "Ising formulations of many NP problems," 2013. arXiv:1302.5843.

[54] A. B. Finnila, M. A. Gomez, C. Sebenik, C. Stenson, and J. D. Doll, "Quantum annealing: A new method for minimizing multidimensional functions," *Chemical Physics Letter*, vol. 219, pp. 343–348, 1994.

[55] T. Kadowaki and H. Nishimori, "Quantum annealing in the transverse Ising model," *Physical Review E*, vol. 58, no. 5, p. 5355, 1998.

[56] E. Farhi, J. Goldstone, S. Gutmann, J. Lapan, A. Lundgren, and D. Preda, "A quantum adiabatic evolution algorithm applied to random instances of an NP-complete problem," *Science*, vol. 292, no. 5516, pp. 472–475, 2001.

[57] A. Das and B. K. Chakrabarti, *Quantum annealing and related optimization methods*, vol. 679. Berlin, Germany: Springer Science & Business Media, 2005.

[58] A. Das and B. K. Chakrabarti, "Quantum annealing and analog quantum computation," *Review of Modern Physics*, vol. 80, p. 1061, 2008.

[59] R. Harris, M. W. Johnson, T. Lanting, A. J. Berkley, J. Johansson, P. Bunyk, E. Tolkacheva, E. Ladizinsky, N. Ladizinsky, T. Oh, F. Cioata, I. Perminov, P. Spear, C. Enderud, C. Rich, S. Uchaikin, M. C. Thom, E. M. Chapple, J. Wang, B. Wilson, M. H. S. Amin, N. Dickson, K. Karimi, B. Macready, C. J. S. Truncik, and G. Rose, "Experimental investigation of an eight-qubit unit cell in a superconducting optimization processor," *Physical Review B*, vol. 82, p. 024511, 2010.

[60] M. W. Johnson, M. H. S. Amin, S. Gildert, T. Lanting, F. Hamze, N. Dickson, R. Harris, A. J. Berkley, J. Johansson, P. Bunyk, E. M. Chapple, C. Enderud, J. P. Hilton, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, C. J. S. Truncik, S. Uchaikin, J. Wang, B. Wilson, and G. Rose, "Quantum annealing with manufactured spins," *Nature*, vol. 473, no. 7346, pp. 194–198, 2011.

[61] V. Bapst, L. Foini, F. Krzakala, G. Somerjian, and F. Zamponi, "The quantum adiabatic algorithm applied to random optimization problems: The quantum spin glass perspective," *Physics Reports*, vol. 523, no. 3, pp. 127–205, 2013.

[62] N. G. Dickson, M. W. Johnson, M. H. Amin, R. Harris, F. Altomare, A. J. Berkley, P. Bunyk, J. Cai, E. M. Chapple, P. Chavez, F. Cioata, T. Cirip, P. deBuen, M. Drew-Brook, C. Enderud, S. Gildert, F. Hamze, J. P. Hilton, E. Hoskinson, K. Karimi, E. Ladizinsky, N. Ladizinsky, T. Lanting, T. Mahon, R. Neufeld, T. Oh, I. Perminov, C. Petroff, A. Przybysz, C. Rich, P. Spear, A. Tcaciuc, M. C. Thom, E. Tolkacheva, S. Uchaikin, J. Wang, A. B. Wilson, Z. Merali, and G. Rose, "Thermally assisted quantum annealing of a 16-qubit problem," *Nature Communications*, vol. 4, p. 1903, 2013.

[63] S. Boixo, T. F. Ronnøw, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, and M. Troyer, "Evidence for quantum annealing with more than one hundred qubits," *Nature Physics*, vol. 10, pp. 218–224, 2014.

[64] C. C. McGeoch and C. Wang, "Experimental evaluation of an adiabiatic quantum system for combinatorial optimization," in *Proceedings of the ACM International Conference on Computing Frontiers*, (New York, USA), 2013.

[65] S. Dash, "A note on QUBO instances defined on Chimera graphs," 2013. arXiv:1306.1202 [math.OC].

[66] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, "Experimental signature of programmable quantum annealing," *Nature Communications*, vol. 4, p. 3067, 2013.

[67] T. Lanting, A. J. Przybysz, A. Y. Smirnov, F. M. Spedalieri, M. H. Amin, A. J. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, N. Dickson, C. Enderud, J. P. Hilton, E. Hoskinson, M. W. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, T. Oh, I. Perminov, C. Rich, M. C. Thom, E. Tolkacheva, S. Uchaikin, A. B. Wilson, and G. Rose, "Entanglement in a quantum annealing processor," *Physical Review X*, vol. 4, p. 021041, 2014.

[68] S. Santra, G. Quiroz, G. V. Steeg, and D. Lidar, "MAX 2-SAT with up to 108 qubits," *New Journal of Physics*, vol. 16, p. 045006, 2014.

[69] T. F. Rønnow, Z. Wang, J. Job, S. Boixo, S. V. Isakov, D. Wecker, J. M. Martinis, D. A. Lidar, and M. Troyer, "Defining and detecting quantum speedup," *Science*, vol. 345, p. 420, 2014.

[70] W. Vinci, K. Markstrom, S. Boixo, A. Roy, F. M. Spedalieri, P. A. Warburton, and S. Severini, "Hearing the shape of the Ising model with a programmable superconducting-flux annealer," *Scientific Report*, vol. 4, p. 5703, 2014.

[71] S. W. Shin, G. Smith, J. A. Smolin, and U. Vazirani, "How "quantum" is the D-Wave machine?," 2014. arXiv:1401.7087 [quant-ph].

[72] C. McGeoch, "Adiabatic quantum computation and quantum annealing: Theory and practice," in *Synthesis Lectures on Quantum Computing*, vol. 5, Morgan & Claypool, 2014.

[73] W. Vinci, T. Albash, G. Paz-Silva, I. Hen, and D. A. Lidar, "Quantum annealing correction with minor embedding," *Physical Review A*, vol. 92, p. 042310, 2015.

[74] T. Albash, T. F. Rønnow, M. Troyer, and D. A. Lidar, "Reexamining classical and quantum models for the D-Wave One processor," *The European Physical Journal Special Topics*, vol. 224, p. 111, 2015.

[75] A. D. King and C. C. McGeoch, "Algorithm engineering for a quantum annealing platform," 2014. arXiv:1410.2628 [cs.DS].

[76] P. J. D. Crowley, T. Duric, W. Vinci, P. A. Warburton, and A. G. Green, "Quantum and classical in adiabatic computation," *Physical Review A*, vol. 90, p. 042317, 2014.

[77] D. S. Steiger, T. F. Rønnow, and M. Troyer, "Heavy tails in the distribution of time-to-solution for classical and quantum annealing," *Physical Review Letters*, vol. 115, p. 230501, 2015.

[78] B. Bauer, L. Wang, I. Pižorn, and M. Troyer, "Entanglement as a resource in adiabatic quantum optimization," 2015. arXiv:1501.06914 [cond-mat.dis-nn].

[79] T. Albash, I. Hen, F. M. Spedalieri, and D. A. Lidar, "Reexamination of the evidence for entanglement in the D-wave processor," *Physical Review A*, vol. 92, p. 062328, 2015.

[80] N. Chancellor, S. Szoke, W. Vinci, G. Aeppli, and P. A. Warburton, "Maximum-entropy inference with a programmable annealer," *Scientific Report*, p. 22318, 2016.

[81] A. Perdomo-Ortiz, B. O'Gorman, J. Fluegemann, R. Biswas, and V. N. Smelyanskiy, "Determination and correction of persistent biases in quantum annealers," 2015. arXiv:1503.05679 [quant-ph].

[82] W. Vinci, T. Albash, and D. A. Lidar, "Nested quantum annealing correction," 2015. arXiv:1511.07084 [quant-ph].

[83] E. Farhi, J. Goldstone, and S. Gutmann, "Quantum adiabatic evolution algorithms versus simulated annealing," *MIT-CTP-3228*, 2002.

[84] G. E. Santoro, R. Martoňák, E. Tosatti, and R. Car, "Theory of quantum annealing of an Ising spin glass," *Science*, vol. 295, pp. 2427–2430, 2002.

[85] B. Heim, T. F. Rønnow, S. V. Isakov, and M. Troyer, "Quantum versus classical annealing of Ising spin glasses," *Science*, vol. 348, no. 6231, pp. 215–217, 2014.

[86] E. Farhi, J. Goldstone, and S. Gutmann, "A numerical study of the performance of a quantum adiabatic evolution algorithm for satisfiability," *MIT-CTP-3006*, 2000.

[87] V. Choi, "Adiabatic quantum algorithms for the NP-complete Maximum-Weight Independent set, Exact Cover and 3SAT problems," 2010. arXiv:1004.2226.

[88] A. M. Childs, E. Farhi, J. Goldstone, and S. Gutmann, "Finding cliques by quantum adiabatic evolution," *Quantum Information and Computation*, vol. 2, no. 181, 2002. MIT-CTP #3067.

[89] X. Peng, Z. Liao, N. Xu, G. Qin, X. Zhou, D. Suter, and J. Du, "Quantum adiabatic algorithm for factorization and its experimental implementation," *Physical Review Letters*, vol. 101, no. 22, p. 220405, 2008.

[90] I. Hen and A. P. Young, "Solving the Graph Isomorphism problem with a quantum annealer," *Physical Review A*, vol. 86, no. 4, p. 042310, 2012.

[91] F. Gaitan and L. Clark, "Graph isomorphism and adiabatic quantum computing," *Physical Review A*, vol. 89, no. 2, p. 022342, 2014.

[92] Z. Bian, F. Chudak, W. G. Macready, L. Clark, and F. Gaitan, "Experimental determination of Ramsey numbers," *Physical Review Letters*, vol. 111, no. 13, p. 130505, 2013.

[93] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready, "Training a binary classifier with the quantum adiabatic algorithm," 2008. arXiv:0811.0416.

[94] V. S. Denchev, N. Ding, S. Vishwanathan, and H. Neven, "Robust classification with adiabatic quantum optimization," 2012. arXiv:1205.1148.

[95] J. Roland and N. J. Cerf, "Quantum search by local adiabatic evolution," *Physical Review A*, vol. 65, p. 042308, 2002.

[96] S. Garnerone, P. Zanardi, and D. A. Lidar, "Adiabatic quantum algorithm for search engine ranking," *Physical Review Letters*, vol. 108, no. 23, p. 230506, 2012.

[97] V. S. Denchev, S. Boixo, S. V. Isakov, N. Ding, R. Babbush, V. Smelyanskiy, J. Martinis, and H. Neven, "What is the computational value of finite range tunneling?," 2015. arXiv:1512.02206.

[98] R. Raussendorf and H. J. Briegel, "A one-way quantum computer," *Physical Review Letter*, vol. 86, pp. 5188–5191, May 2001.

[99] R. Raussendorf, D. E. Browne, and H. J. Briegel, "Measurement-based quantum computation on cluster states," *Physical Review A*, vol. 68, p. 022312, Aug 2003.

[100] H. J. Briegel and R. Raussendorf, "Persistent entanglement in arrays of interacting particles," *Physical Review Letter*, vol. 86, pp. 910–913, Jan 2001.

[101] R. Raussendorf, S. Bravyi, and J. Harrington, "Long-range quantum entanglement in noisy cluster states," *Physical Review A*, vol. 71, p. 062313, Jun 2005.

[102] D. A. Meyer, "From quantum cellular automata to quantum lattice gases," *Journal of Statistical Physics*, vol. 85, no. 5, pp. 551–574, 1996.

[103] S. Wiesner, "Simulations of many-body quantum systems by a quantum computer," 1996. arXiv:quant-ph/9603028.

[104] D. S. Abrams and S. Lloyd, "Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors," *Physical Review Letter*, vol. 83, pp. 5162–5165, Dec 1999.

[105] D. A. Lidar and O. Biham, "Simulating Ising spin glasses on a quantum computer," *Physical Review E*, vol. 56, pp. 3661–3681, Sep 1997.

[106] B. M. Boghosian and W. Taylor, "Proceedings of the fourth workshop on physics and consumption simulating quantum mechanics on a quantum computer," *Physica D: Nonlinear Phenomena*, vol. 120, no. 1, pp. 30 – 42, 1998.

[107] C. Zalka, "Simulating quantum systems on a quantum computer," *Proceedings of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, vol. 454, no. 1969, pp. 313–322, 1998.

[108] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders, "Efficient quantum algorithms for simulating sparse hamiltonians," *Communications in Mathematical Physics*, vol. 270, no. 2, pp. 359–371, 2007.

[109] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, "Polynomial-time quantum algorithm for the simulation of chemical dynamics," *Proceedings of National Academy of Science*, vol. 105, no. 48, pp. 18681–18686, 2008.

[110] N. Wiebe, D. Berry, P. Hyer, and B. C. Sanders, "Higher order decompositions of ordered operator exponentials," *Journal of Physics A: Mathematical and Theoretical*, vol. 43, no. 6, p. 065203, 2010.

[111] N. J. Ward, I. Kassal, and A. Aspuru-Guzik, "Preparation of many-body states for quantum simulation," *The Journal of Chemical Physics*, vol. 130, no. 19, 2009.

[112] S. Raeisi, N. Wiebe, and B. C. Sanders, "Quantum-circuit design for efficient simulations of many-body quantum dynamics," *New Journal of Physics*, vol. 14, no. 10, p. 103017, 2012.

[113] B. C. Sanders, *Efficient Algorithms for Universal Quantum Simulation*, pp. 1–10. Berlin / Heidelberg, Germany: Springer Berlin / Heidelberg, 2013.

[114] W. Pauli, "The connection between spin and statistics," *Physical Review*, vol. 58, pp. 716–722, Oct 1940.

[115] P. van Nieuwenhuizen and A. Waldron, "On Euclidean spinors and Wick rotations," *Physics Letters B*, vol. 389, no. 1, pp. 29 – 36, 1996.

[116] J. Schwinger, "The theory of quantized fields. I," *Physical Review*, vol. 82, pp. 914–927, Jun 1951.

[117] E. U. Condon and G. H. Shortley, *The Theory of Atomic Spectra*. Cambridge University Press, 1967.

[118] P. Jordan and E. Wigner, "Über das Paulische äquivalenzverbot," *Zeitschrift für Physik*, vol. 47, no. 9-10, pp. 631–651, 1928.

[119] J. T. Seeley, M. J. Richard, and P. J. Love, "The Bravyi-Kitaev transformation for quantum computation of electronic structure," *The Journal of Chemical Physics*, vol. 137, no. 22, 2012.

[120] S. B. Bravyi and A. Y. Kitaev, "Fermionic quantum computation," *Annals of Physics*, vol. 298, no. 1, pp. 210 – 226, 2002.

[121] R. Babbush, P. J. Love, and A. Aspuru-Guzik, "Adiabatic quantum simulation of quantum chemistry," *Scientific Reports*, vol. 4, no. 6603, 2014. arXiv:1311.3967 [quant-ph].

[122] A. Y. Kitaev, "Fault-tolerant quantum computation by anyons," *Annals of Physics*, vol. 303, no. 1, pp. 2 – 30, 2003.

[123] P. Fendley, "Quantum loop models and the non-abelian toric code," 2007. arXiv:0711.0014 [cond-mat.stat-mech].

[124] J. Kempe and O. Regev, "3-Local Hamiltonian is QMA-complete," *Quantum Information and Computation*, vol. 3, no. 3, pp. 258–64, 2003.

[125] D. Nagaj and S. Mozes, "New construction for a QMA complete 3-local Hamiltonian," *Journal of Mathematical Physics*, vol. 48, no. 7, 2007.

[126] M. H. Sergey Bravyi, "On complexity of the quantum Ising model," 2014. arXiv:1410.0703 [quant-ph].

[127] J. D. Biamonte, V. Bergholm, J. D. Whitfield, J. Fitzsimons, and A. Aspuru-Guzik, "Adiabatic quantum simulators," *AIP Advances*, vol. 1, no. 2, p. 022126, 2011.

[128] L. Veis and J. Pittner, "Adiabatic state preparation study of methylene," *ArXiv e-prints*, Jan. 2014. arXiv:1401.3186 [quant-ph].

[129] S. Bravyi, D. DiVincenzo, D. Loss, and B. Terhal, "Quantum simulation of many-body Hamiltonians using perturbation theory with bounded-strength interactions," *Physical Review Letter*, vol. 101, p. 070503, 2008.

[130] S. Bravyi, D. P. DiVincenzo, R. I. Oliveira, and B. M. Terhal, "The complexity of stoquastic Local Hamiltonian problems," *Quantum Information and Computation*, vol. 8, no. 5, pp. 0361–0385, 2008.

[131] A. Ganti, U. Onunkwo, and K. Young, "A family of [[6k, 2k, 2]] codes for practical, scalable adiabatic quantum computation," Sept. 2013. arXiv:1309.1674 [quant-ph].

[132] C. Bloch, "Sur la théorie des perturbations des états liés," *Nuclear Physics*, vol. 6, pp. 329–347, 1958.

[133] M. D. Price, S. S. Somaroo, A. E. Dunlop, T. F. Havel, and D. G. Cory, "Generalized methods for the development of quantum logic gates for an NMR quantum information processor," *Physical Review A*, vol. 60, pp. 2777–2780, 1999.

[134] C. H. Tseng, S. S. Somaroo, Y. S. Sharf, E. Knill, R. Laflamme, T. F. Havel, and D. G. Cory, "Quantum simulation of a three-body interaction Hamiltonian on an NMR quantum computer," *Physical Review A*, vol. 61, pp. 12302–12308, 2000. arXiv:quant-ph/9908012.

[135] R. Harris, A. J. Berkley, M. W. Johnson, P. Bunyk, S. Govorkov, M. C. Thom, S. Uchaikin, A. B. Wilson, J. Chung, E. Holtham, J. D. Biamonte, A. Y. Smirnov, M. H. S. Amin, and A. Maassen van den Brink, "Sign- and magnitude-tunable coupler for superconducting flux qubits," *Physical Review Letter*, vol. 98, p. 177001, Apr 2007.

[136] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar, "Experimental signature of programmable quantum annealing," *Nature Communications*, vol. 4, p. 2067, June 2012. arXiv:1212.1739 [quant-ph].

[137] K. L. Pudenz, T. Albash, and D. A. Lidar, "Error corrected quantum annealing with hundreds of qubits," *Nature Communications*, vol. 5, p. 3243, 2014. arXiv:1307.8190 [quant-ph].

[138] J. D. Biamonte, "Non-perturbative k-body to two-body commuting conversion hamiltonians and embedding problem instances into Ising spins," *Physical Review A*, vol. 77, no. 5, p. 052331, 2008. arXiv:0801.3800.

[139] J. D. Whitfield, M. Faccin, and J. D. Biamonte, "Ground state spin logic," *European Physics Letter*, vol. 99, no. 57004, 2012. arXiv:1205.1742v1.

[140] R. Babbush, B. O'Gorman, and A. Aspuru-Guzik, "Resource efficient gadgets for compiling adiabatic quantum optimization problems," *Annals of Physics*, vol. 525, no. 10-11, pp. 877–888, 2013. arXiv:1307.8041 [quant-ph].

[141] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*. New York, NY, USA: W. H. Freeman & Co., 1979.

[142] I. Dinur, "The PCP theorem by gap amplification," *Journal of ACM*, vol. 54, June 2007.

[143] D. Aharonov, I. Arad, and T. Vidick, "Guest column: The quantum PCP conjecture," *SIGACT News*, vol. 44, pp. 47–79, June 2013.

[144] M. B. Hastings, "An area law for one-dimensional quantum systems," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2007, no. 08, p. P08024, 2007.

[145] D. Aharonov, I. Arad, Z. Landau, and U. Vazirani, "The 1D area law and the complexity of quantum states: A combinatorial approach," in *IEEE 52nd Annual Symposium on Foundations of Computer Science (FOCS)*, pp. 324–333, Oct 2011.

[146] J. Eisert, M. Cramer, and M. B. Plenio, "Colloquium: Area laws for the entanglement entropy," *Review of Modern Physics*, vol. 82, pp. 277–306, Feb 2010.

[147] R. Koenig, "Simplifying quantum double Hamiltonians using perturbative gadgets," *Quantum Information and Computation*, vol. 10, no. 3, pp. 292–324, 2010. arXiv:0901.1333 [quant-ph].

[148] S. A. Ocko and B. Yoshida, "Nonperturbative gadget for topological quantum codes," *Physical Review Letter*, vol. 107, no. 250502, 2011. arXiv:1107.2697 [quant-ph].

[149] S. Bravyi, D. DiVincenzo, and D. Loss, "Schrieffer-Wolff transformation for quantum many-body systems," *Annals of Physics*, vol. 326, no. 10, 2011. arXiv:1105.0675.

[150] K. C. Young, R. Blume-Kohout, and D. A. Lidar, "Adiabatic quantum optimization with the wrong Hamiltonian," *Physical Review A*, vol. 88, p. 062314, Dec 2013.

[151] D. Janzing and P. Wocjan, "Ergodic quantum computing," *Quantum Information Processing*, vol. 4, pp. 129–158, June 2005.

[152] D. Nagaj, "Universal two-body-Hamiltonian quantum computing," *Physical Review A*, vol. 85, p. 032330, Mar 2012.

[153] M. Walschaers, J. F. de Cossio Diaz, R. Mulet, and A. Buchleitner, "Optimally designed quantum transport across disordered networks," *Physical Review Letter*, vol. 111, p. 180601, Oct 2013.

[154] C. M. Bender and S. A. Orszag, *Advanced Mathematical Methods for Scientists and Engineers I: Asympotic Methods and Perturbation Theory*. New York, USA: Springer Science & Business Media, 1999.

[155] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry*. New York, USA: Dover Publications, 1996.

[156] H. Primas, "Generalized perturbation theory in operator form," *Review of Modern Physics*, vol. 35, no. 710, 1963.

[157] D. R. Herschbach, J. S. Avery, and O. Goscinski, *Dimensional scaling in chemical physics*. Dordrecht, Netherlands: Springer Science & Business Media, 1993.

[158] M. J. Martin, M. Bishof, M. D. Swallows, X. Zhang, C. Benko, J. von Stecher, A. V. Gorshkov, A. M. Rey, and J. Ye, "A quantum many-body spin system in an optical lattice clock," *Science*, vol. 341, no. 6146, pp. 632–636, 2013.

[159] R. Islam, C. Senko, W. C. Campbell, S. Korenblit, J. Smith, A. Lee, E. E. Edwards, C.-C. J. Wang, J. K. Freericks, and C. Monroe, "Emergence and frustration of magnetism with variable-range interactions in a quantum simulator," *Science*, vol. 340, no. 6132, pp. 583–587, 2013.

[160] I. G. MacDonald, *Symmetric Functions and Hall Polynomials*. Oxford, UK: Oxford University Press, 1998.

[161] J. von Neumann, "The general and logical theory of automata," in *Design of Computers, Theory of Automata and Numerical Analysis* (A. H. Taub, ed.), vol. V, Pergamon Press, 1963.

[162] N. Wiener and A. Rosenblueth, "The mathematical formulation of the problem of conduction of connected excitable elements, specifically in cardiac muscle," *Archivos del Instituto de Cardiología de México*, vol. 16, no. 3, pp. 205–265, 1946.

[163] M. Gardner, "The fantastic combinations of John Conway's new solitaire game "life"," *Scientific American*, vol. 223, pp. 120–123, 1970.

[164] A. R. Smith, "Simple non-trivial self-reproducing machines," in *Artificial Life II, SFI studies in the Sciences of Complexity* (C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, eds.), vol. X, Addison-Wesley, 1991.

[165] C. E. Shannon, "Von Neumann's contributions to automata theory," *Bulletin of American Mathematics Society*, vol. 64, pp. 123–129, 1958.

[166] S. Wolfram, "Statistical mechanics of cellular automata," *Review of Modern Physics*, vol. 55, no. 3, pp. 601–644, 1983.

[167] T. Toffoli and N. Margolus, *Cellular Automata Machines: A New Environment for Modeling*. Cambridge, MA, USA: MIT Press, 1987.

[168] E. N. K. Eloranta, "The kink of cellular automaton rule 18 performs a random walk," *Journal of Statistical Physics*, vol. 69, pp. 1131–1136, 1992.

[169] K. Eloranta, "Random walks in cellular automata," *Nonlinearity*, vol. 6, no. 6, p. 1025, 1993.

[170] F. Verstratete, J. I. Cirac, and V. Murg, "Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems," *Advances in Physics*, vol. 57, no. 143, 2008.

[171] J. I. Cirac and F. Verstraete, "Renormalization and tensor product states in spin chains and lattices," *Journal of Physics A: Mathematical and Theoretical*, vol. 42, no. 504004, 2009.

[172] R. Augusiak, F. M. Cucchietti, and M. Lewenstein, "Many-body physics from a quantum information perspective," in *Lecture Notes in Physics* (D. C. Cabra, A. Honecker, and P. Pujol, eds.), vol. 843, ch. 6, pp. 245–294, Berlin / Heidelberg, Germany: Springer-Verlag, 2012.

[173] S. Lloyd and B. Terhal, "Adiabatic and Hamiltonian computing on a 2D lattice with simple 2-qubit interactions," *New Journal of Physics*, vol. 18, p. 023042, 2016.

[174] R. Rucker, "Continuous-valued cellular automata in two dimensions," in *New Constructions in Cellular Automata* (D. Griffeath and C. Moore, eds.), Santa Fe Institute Studies in the Sciences of Complexity Proceedings, (Santa Fe, NM), 1999.

[175] F. Jiménez-Morales, "Evolving three-dimensional cellular automata to perform a quasiperiod-3 collective behavior task," *Physical Review E*, vol. 60, no. 4934, 1999.

[176] S. Wolfram, "Twenty problems in the theory of cellular automata," *Physica Scripta*, vol. 1985, no. T9, 1985.

[177] S. N. Cole, "Real-time computation by n-dimensional iterative arrays of finite-state machines," *IEEE Transactions on Computers*, vol. 18, no. 4, pp. 349–365, 1969.

[178] Y. Cao, "Algorithm for evaluating any n-variable constant-degree symmetric polynomial in $O(n)$ time," 2016. http://web.ics.purdue.edu/~cao23/computing-symmetric-polynomials.pdf.

APPENDICES

A. COMPENSATION FOR THE 4-LOCAL ERROR TERMS IN PARALLEL 3- TO 2-BODY GADGET

Continuing the discussion in Section 2.5, here we deal with $\Theta(1)$ error terms that arise in the 3rd and 4th order perturbative expansion when $\mathbf{V}$ in Eq. 2.52 is without $\mathbf{V}_3$ and in so doing explain the construction of $\bar{\mathbf{V}}_{ij}$ in Eq. 2.59. From the previous description of the 3rd and 4th order terms, for each pair of terms (i) and (j) where i and j are integers between 1 and m, let

$$\mathbf{M}_1 = (\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i)(\kappa_j \mathbf{A}_j + \lambda_j \mathbf{B}_j) \tag{A.1}$$

$$\mathbf{M}_2 = (\kappa_j \mathbf{A}_j + \lambda_j \mathbf{B}_j)(\kappa_i \mathbf{A}_i + \lambda_i \mathbf{B}_i) \tag{A.2}$$

and then the $\Theta(1)$ error term arising from the 3rd and 4th order perturbative expansion can be written as

$$\frac{1}{(z-\Delta)^2}\left[\frac{1}{z-2\Delta}(\mathbf{M}_1^2 + \mathbf{M}_2^2) + \left(\frac{1}{\Delta} + \frac{1}{z-2\Delta}\right)(\mathbf{M}_1\mathbf{M}_2 + \mathbf{M}_2\mathbf{M}_1)\right]. \tag{A.3}$$

Based on the number of non-commuting pairs among $\mathbf{A}_i$, $\mathbf{A}_j$, $\mathbf{B}_i$ and $\mathbf{B}_j$, all possible cases can be enumerated as the following:

case 0:		$[\mathbf{A}_i, \mathbf{A}_j] = 0, [\mathbf{B}_i, \mathbf{B}_j] = 0, [\mathbf{A}_i, \mathbf{B}_j] = 0, [\mathbf{B}_i, \mathbf{A}_j] = 0$
case 1:	1.1 :	$[\mathbf{A}_i, \mathbf{A}_j] = 0, [\mathbf{B}_i, \mathbf{B}_j] = 0, [\mathbf{A}_j, \mathbf{B}_i] \neq 0$
	1.2 :	$[\mathbf{A}_i, \mathbf{A}_j] = 0, [\mathbf{B}_i, \mathbf{B}_j] = 0, [\mathbf{A}_i, \mathbf{B}_j] \neq 0$
	1.3 :	$[\mathbf{A}_i, \mathbf{A}_j] = 0, [\mathbf{B}_i, \mathbf{B}_j] \neq 0$
	1.4 :	$[\mathbf{A}_i, \mathbf{A}_j] \neq 0, [\mathbf{B}_i, \mathbf{B}_j] = 0$
case 2:		$[\mathbf{A}_i, \mathbf{A}_j] \neq 0, [\mathbf{B}_i, \mathbf{B}_j] \neq 0.$

(A.4)

In case 0, clearly $\mathbf{M}_1 = \mathbf{M}_2$. Then the $\Theta(1)$ error becomes

$$\frac{1}{(z-\Delta)^2}\left(\frac{1}{\Delta} + \frac{2}{z-2\Delta}\right) \cdot 2\mathbf{M}_1^2 = \Theta(\Delta^{-1})$$

which does not need any compensation. In case 1, for example in the subcase 1.1, $\mathbf{A}_j$ does not commute with $\mathbf{B}_i$. Then $\mathbf{M}_1$ and $\mathbf{M}_2$ can be written as

$$\mathbf{M}_1 = \mathbf{K} + \kappa_j \lambda_i \mathbf{B}_i \mathbf{A}_j \tag{A.5}$$

$$\mathbf{M}_2 = \mathbf{K} + \kappa_j \lambda_i \mathbf{A}_j \mathbf{B}_i \tag{A.6}$$

where $\mathbf{K}$ contains the rest of the terms in $\mathbf{M}_1$ and $\mathbf{M}_2$. Furthermore,

$$\mathbf{M}_1^2 + \mathbf{M}_2^2 = 2\mathbf{K}^2 - 2(\kappa_j \lambda_i)^2 \mathbf{I} \tag{A.7}$$

$$\mathbf{M}_1 \mathbf{M}_2 + \mathbf{M}_2 \mathbf{M}_1 = 2\mathbf{K}^2 + 2(\kappa_j \lambda_i)^2 \mathbf{I}. \tag{A.8}$$

Hence the $\Theta(1)$ term in this case becomes

$$\frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta} + \frac{2}{z-2\Delta}\right) 2\mathbf{K}^2 + \frac{1}{\Delta} \cdot 2(\kappa_j \lambda_i)^2 \mathbf{I}\right] \tag{A.9}$$

where the first term is $\Theta(\Delta^{-1})$ and the second term is $\Theta(1)$, which needs to be compensated. Similar calculations for cases 1.2, 1.3 and 1.4 will yield $\Theta(1)$ error with the same norm. In case 2, define $\mathbf{R} = \kappa_i \lambda_j \mathbf{A}_i \mathbf{B}_j + \lambda_i \kappa_j \mathbf{B}_i \mathbf{A}_j$ and $\mathbf{T} = \kappa_i \kappa_j \mathbf{A}_i \mathbf{A}_i + \lambda_i \lambda_j \mathbf{B}_i \mathbf{B}_i$. Then

$$\mathbf{M}_1^2 + \mathbf{M}_2^2 = 2(\mathbf{R}^2 + \mathbf{T}^2) \tag{A.10}$$

$$\mathbf{M}_1 \mathbf{M}_2 + \mathbf{M}_2 \mathbf{M}_1 = 2(\mathbf{R}^2 - \mathbf{T}^2). \tag{A.11}$$

The $\Theta(1)$ error terms in the 3rd and 4th order perturbative expansion becomes

$$\frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta} + \frac{2}{z-2\Delta}\right) \cdot 2\mathbf{R}^2 - \frac{1}{\Delta} \cdot 2\mathbf{T}^2\right] \tag{A.12}$$

where the first term is $\Theta(\Delta^{-1})$ and hence needs no compensation. The second term is $\Theta(1)$. Define

$$s_0^{(i,j)} = \begin{cases} 1 & \text{if case 0} \\ 0 & \text{Otherwise} \end{cases} \tag{A.13}$$

With the definitions of $s_1^{(i,j)}$ and $s_2^{(i,j)}$ in Eq. 2.56, Eq. 2.57 and Eq. 2.58, the contribution of the i-th and the j-th target terms to the $\Theta(1)$ error in the perturbative expansion $\mathbf{\Sigma}_-(z)$ becomes

$$\begin{aligned} & s_0^{(i,j)} \cdot \frac{1}{(z-\Delta)^2}\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2(\kappa_i\mathbf{A}_i+\lambda_i\mathbf{B}_i)^2(\kappa_j\mathbf{A}_j+\lambda_j\mathbf{B}_j)^2 \\ + \quad & s_1^{(i,j)} \cdot \frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2\mathbf{K}_{ij}^2+\frac{1}{\Delta}\cdot 2(\kappa_i\kappa_j)^2\mathbf{I}\right] \\ + \quad & s_2^{(i,j)} \cdot \frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2\mathbf{R}_{ij}^2+\frac{1}{\Delta}\cdot 2\{[(\kappa_i\kappa_j)^2+(\lambda_i\lambda_j)^2]\mathbf{I}\right. \\ & \left. -2\kappa_i\kappa_j\lambda_i\lambda_j\mathbf{A}_i\mathbf{A}_j\mathbf{B}_i\mathbf{B}_j\}\right]. \end{aligned} \tag{A.14}$$

The term proportional to $s_0^{(i,j)}$ in Eq. A.14 does not need compensation since it is already $\Theta(\Delta^{-1})$. The term proportional to $s_1^{(i,j)}$ can be compensated by the corresponding term in $\bar{\mathbf{V}}_{ij}$ in Eq. 2.59 that is proportional to $s_1^{(i,j)}$. Similarly, the $\Theta(1)$ error term proportional to $s_2^{(i,j)}$ can be compensated by the term in $\bar{\mathbf{V}}_{ij}$ in Eq. 2.59 that is proportional to $s_2^{(i,j)}$.

Now we deal with generating the 4-local term in $\bar{\mathbf{V}}_{ij}$. Introduce an ancilla u_{ij} and construct a gadget $\tilde{\mathbf{H}}_{ij}=\mathbf{H}_{ij}+\mathbf{V}_{ij}$ such that $\mathbf{H}_{ij}=\Delta|1\rangle\langle 1|_{u_{ij}}$ and the perturbation $\mathbf{V}_{ij}$ becomes

$$\mathbf{V}_{ij}=(\kappa_i\mathbf{A}_i+\lambda_j\mathbf{B}_j)\otimes\mathbf{X}_{u_{ij}}+(\kappa_j\mathbf{A}_j+\lambda_i\mathbf{B}_i)\otimes|1\rangle\langle 1|_{u_{ij}}+\mathbf{V}'_{ij} \tag{A.15}$$

where $\mathbf{V}'_{ij}$ is defined as

$$\mathbf{V}'_{ij}=\frac{1}{\Delta}(\kappa_i\mathbf{A}_i+\lambda_j\mathbf{B}_j)^2+\frac{1}{\Delta^3}\left[(\kappa_j^2+\lambda_i^2)(\kappa_i\mathbf{A}_i+\lambda_j\mathbf{B}_j)^2-2\kappa_j\lambda_i(\kappa_j^2+\lambda_j^2)\mathbf{A}_j\mathbf{B}_i\right] \tag{A.16}$$

The self-energy expansion $\mathbf{\Sigma}_-(z)$ is now

$$\mathbf{\Sigma}_-(z)=\frac{1}{(z-\Delta)^3}4\kappa_i\kappa_j\lambda_i\lambda_j\mathbf{A}_i\mathbf{A}_j\mathbf{B}_i\mathbf{B}_j+O(\Delta^{-1/2})$$

which is $O(\Delta^{-1/2})$ close to the 4-local compensation term in $\bar{\mathbf{V}}_{ij}$. We apply the the gadget $\tilde{\mathbf{H}}_{ij}$ for every pair of qubits with $s_2^{(i,j)}=1$. The cross-gadget contribution between the $\tilde{\mathbf{H}}_{ij}$ gadgets as well as those cross-gadget contribution between $\tilde{\mathbf{H}}_{ij}$ gadgets

and gadgets based on ancilla qubits u_1 through u_m both belong to the case 1 of the Eq. A.4 and hence are easy to deal with using 2-body terms.

B. UPPER BOUNDS ON LOW-ORDER PERTURBATION SERIES TERMS FOR 2-BODY GADGETS

In this Appendix, for the purpose of illustration we calculate upper bounds on the norm of the first few orders in the perturbation series for the self-energy for our 2-body gadget construction from Section 3.3.2.

The 2^{nd} order. This order is what contributes to the effective Hamiltonian, which has M terms of norm $O(1)$ there. Let us see what we get here. From (3.48) we see that $\mathbf{T}_2 = \frac{1}{z-\Delta}\sum_{j=1}^{M} R\beta_j^2(\mathbf{A}_{a_j} - \mathbf{B}_{b_j})^2$. Every term at the second order corresponds to a transition of the form

$$\mathcal{L}_- \to |y\rangle \to \mathcal{L}_-. \tag{B.1}$$

Here $|y\rangle$ is a state where only one direct ancilla qubit w is flipped to $|1\rangle$ while the others remain at $|0\rangle$. From our construction of V in (3.13), observe that each term that involves a particular direct ancilla $w_i^{(j)}$ is associated with a corresponding coefficient β_j. Therefore all the transitions of the form (B.1) involving $w_i^{(j)}$ would contribute a term of the form

$$\underbrace{\beta_j(\mathbf{A}_{a_j} - \mathbf{B}_{b_j})}_{\mathbf{V}_{-+}} \cdot \underbrace{\frac{1}{z-\Delta}}_{\mathbf{G}_+} \cdot \underbrace{\beta_j(\mathbf{A}_{a_j} - \mathbf{B}_{b_j})}_{\mathbf{V}_{+-}} \tag{B.2}$$

to the perturbative expansion $\mathbf{\Sigma}_-(z)$. Note that because the Hamming weight of y is $h(y) = 1$, the resolvent component $\mathbf{G}_+$ contributes a factor $\frac{1}{z-h(y)\Delta} = \frac{1}{z-\Delta}$. Since R direct ancillas are introduced for the target 2-local term involving a_j and b_j, the total contribution of the direct ancillas used for generating the j-th target term would be multiplied by a factor of R. Summing over all the target terms from $j = 1$ to M, we get the current form of T_2. Assuming $\mathbf{A}_{a_j}$ and $\mathbf{B}_{b_j}$ are both unit-norm operators,

$$\|T_2\| \leq \frac{1}{\Delta} \cdot MR(2\beta_{\max})^2 = 2M\gamma_{\max}, \tag{B.3}$$

using the choice $\beta_i = \sqrt{\frac{\gamma_i \Delta}{2R}}$. This is just what we expected (because the norm of what we are generating should be something on the order of M).

The 4^{th} order. Transitions at the 4^{th} order could involve one or two direct ancillas[1]. In the former case the transition would take the form of

$$\mathcal{L}_- \to |y\rangle \to |y\rangle \to |y\rangle \to \mathcal{L}_- \tag{B.4}$$

where y is a string of Hamming weight 1. Such processes all contribute 0 to the perturbative expansion since $\|\mathbf{H}_{\text{else}}\| = 0$. Now we consider processes that involve two different direct ancillas w_a and w_b. There are two possibilities:

$$\uparrow_a \uparrow_b \downarrow_a \downarrow_b, \qquad \uparrow_a \uparrow_b \downarrow_b \downarrow_a \tag{B.5}$$

where $\uparrow_a$ means flipping w_a from $|0\rangle$ to $|1\rangle$ and $\downarrow_a$ from $|1\rangle$ to $|0\rangle$. Similar for w_b. From $N = MR$ direct ancillas, there are in total $N(N-1)$ ways to choose w_a and w_b. For a fixed choice of w_a and w_b, each of the possible transitions listed above gives rise to at most $(2\beta_{\max})^4$ from the 4 flipping processes (from the above discussion on (B.2) each flipping process contributes a factor of $2\beta_j \leq 2\beta_{\max}$ in $\|\mathbf{T}_k\|$). The $\mathbf{G}_+$ terms contribute an overall factor of $\frac{1}{z-\Delta} \cdot \frac{1}{z-2\Delta} \cdot \frac{1}{z-\Delta}$ to the perturbative expansion. In particular the factor 2 in the component $\frac{1}{z-2\Delta}$ is due to the fact that after the second flipping process the state has two ancillas flipped to 1, resulting in a state $|y'\rangle$ with $h(y') = 2$. Combining these arguments, we have

$$\begin{aligned} \|\mathbf{T}_4\| &\leq 2N(N-1) \cdot (2\beta_{\max})^4 \cdot \frac{1}{\Delta \cdot (2\Delta) \cdot \Delta} = \frac{N(N-1)(2\beta_{\max})^4}{\Delta^3} \\ &\leq \left(\frac{N(2\beta_{\max})^2}{\Delta}\right)^2 \frac{1}{\Delta} = 2M\gamma_{\max} \cdot \left(\frac{2M\gamma_{\max}}{\Delta}\right). \end{aligned} \tag{B.6}$$

Note that compared with the 2^{nd} order term, we collect a factor of $2M\gamma_{\max}/\Delta$ in the upper bound for $\|\mathbf{T}_4\|$.

[1]See also [1] for a detailed explanation.

6^{th} *order.* Following the same notation as before, at 6^{th} order the following transitions contribute non-trivially to $\|\mathbf{T}_6\|$:

$$\uparrow_a\uparrow_b\uparrow_c(\downarrow)^3, \qquad \uparrow_a\uparrow_b\downarrow_a\uparrow_c(\downarrow)^2. \tag{B.7}$$

The former type of transitions has $N(N-1)(N-2)\cdot 6$ different ways of occuring and the $\mathbf{G}_+$ terms contribute a factor of $\frac{1}{z-\Delta}\cdot\frac{1}{z-2\Delta}\cdot\frac{1}{z-3\Delta}\cdot\frac{1}{z-2\Delta}\cdot\frac{1}{z-\Delta}$. The latter has $N(N-1)\cdot 2\cdot(N-1)\cdot 2$ different ways of occuring and a factor $\frac{1}{z-\Delta}\cdot\frac{1}{z-2\Delta}\cdot\frac{1}{z-\Delta}\cdot\frac{1}{z-2\Delta}\cdot\frac{1}{z-\Delta}$ from the G_+ components. Both types involve 6 flipping processes, which amounts to a factor of $(2\beta_{\max})^6$. Hence

$$\|\mathbf{T}_6\| = 6N(N-1)(N-2)(2\beta_{\max})^6\frac{1}{\Delta^2(2\Delta)^2(3\Delta)} \tag{B.8}$$

$$+\,4N(N-1)(N-1)(2\beta_{\max})^6\frac{1}{\Delta^3(2\Delta)^2} \tag{B.9}$$

$$\leq \frac{3}{2}\left(\frac{N(2\beta_{\max})^2}{\Delta}\right)^3\frac{1}{\Delta^2} = 3M\gamma_{\max}\left(\frac{2M\gamma_{\max}}{\Delta}\right)^2. \tag{B.10}$$

Note that another $2M\gamma_{\max}/\Delta$ factor is collected at the 6^{th} order compared with the 4^{th}. Given our choice that $\Delta = M^3R^d$, it is clear that $2M\gamma_{\max}/\Delta = O(M^{-2}R^{-d})$. It is reasonable to speculate that $\|\mathbf{T}_{2m}\| = O(M^{-2m}R^{-dm})$ converges exponentially as $m\to\infty$, which implies that the series $\sum_{m=2}^{\infty}\|\mathbf{T}_{2m}\|$ converges.

C. GLOSSARY OF NOTATIONS FOR CHAPTER 4

As a general guideline, throughout Chapter 4 we use lower case Greek letters for scalar quantities, lower case bold English letters for representing vectors and capital case English letters for representing matrices and operators. Calligraphic fonts (such as $\mathcal{H}$ for the letter 'H') are reserved for representing vector spaces and sets of vertices (as in $\mathcal{E}$). For a vector $\mathbf{v}$, the subscript in the notation v_i represents the i-th element of $\mathbf{v}$. Superscripts in parentheses have two possible meanings: depending on the context, they could mean either the subsystem that the operator acts on (as in Figure 1a of the main text) or the step in a walk. Tables C.1 and C.2 contain the main recurring notations introduced in Chapter 4.

Table C.1: Table of notations (English alphabet) that have recurring appearances in Chapter 4.

Symbol	Meaning and first appearance
$\mathbf{a}$	Partition of a symmetric polynomial $m_{\mathbf{a}}$, see Definition 4.2.1 in Section 4.2.4
$\mathbf{b}$	Reduced partition of a monomial symmetric polynomial. See Definition 4.2.2.
$\mathcal{B}$	Hilbert space for the "bath" in the basic setting in Figure 1a of the main text.
$\mathbf{B}^{(i)}_{pq,jk}$	The pq-th block of $\mathbf{O}^{(i)}_{jk}$ (Eq. 4.21). It contributes a term $\mathbf{B}^{(i)}_{pq,jk} \otimes \lvert\psi^{(i)}_{j,p}\rangle\langle\psi^{(i)}_{k,q}\rvert$ to $\mathbf{V}$. See Eq. 4.25.

$\mathbf{c}$, $\mathbf{c}(\|\psi\rangle)$	Energy configuration of an eigenstate $\|\psi\rangle$ of $\mathbf{H}$. See Def. 4.3.4.
$\tilde{\mathbf{c}}$, $\tilde{\mathbf{c}}(\mathbf{c})$	Reduced energy configuration of a set of $\mathbf{H}$ eigenstates with energy configuration $\mathbf{c}$. See Def. 4.3.9.
$\hat{\mathbf{c}}(\mathbf{n})$	Uniquely reduced configuration associated with an energy combination $\mathbf{n}$. See Def. 4.4.1.
$E_i^{(j)}$	The i-th energy level of the subsystem $\mathbf{H}^{(j)}$ (Fig. 2b of the main text). Also written as E_i.
$E^{(i)}$	The energy of the i-th step during a walk in $\mathbf{H}$ eigenstates, $\mathbf{c}$, $\tilde{\mathbf{c}}$ or $\mathbf{n}$. See Def. 4.3.6.
$E(\mathbf{n})$	The energy of an energy combination $\mathbf{n}$. See Equation 4.26.
$\mathbf{G}(z)$	Operator-valued resolvent, or Green's function. See Section 4.2.1 after Equation 4.8.
$G(\mathcal{V}, \mathcal{E})$	The graph generated by Algorithm 2. $\mathcal{V}$ and $\mathcal{E}$ are the sets of nodes and edges respectively.
$\mathcal{H}^{(i)}$	Hilbert space of the i-th subsystem, see text after Equation 4.9.
$\mathbf{H}$	Unperturbed Hamiltonian for all subsystems (Figure 1a of the main text)
$\mathbf{H}^{(i)}$	The Hamiltonian for the i-th unperturbed subsystem. See Equation 4.11.
$\mathbf{H}_{\mathcal{B}}$	The part of $\tilde{\mathbf{H}}$ that only acts on $\mathcal{B}$. See Equation 4.8.
$\tilde{\mathbf{H}}$	Perturbed Hamiltonian that equals to $\mathbf{H} + \mathbf{V}$. See Section 4.2.1 after Equation 4.8.

ℓ	Total number of energy levels in each subsystem $\mathbf{H}^{(i)}$. See Section 4.2.1 after Equation 4.8.
$\mathcal{L}_-$, $\mathcal{L}_+$	Low- and high- energy subspaces of $\mathbf{H}$. See Section 4.2.1 after Equation 4.8.
$\mathcal{L}_-^{(i)}$, $\mathcal{L}_+^{(i)}$	The low- and high- energy subspace of $\mathbf{H}^{(i)}$.
m	Total number of subsystems. See Figure 1a of the main text and Equation 4.11.
$m_{\mathbf{b}}(\mathbf{x})$	Symmetric polynomial over variables $\mathbf{x} \in \mathbb{C}^n$ with reduced partition $\mathbf{b}$. See Section 4.2.4.
$\mathbf{M}$, M_{jk}	Basic quantity for constructing an upper bound to $\|\mathbf{T}_r\|_2$. See Definition 4.3.3.
$\mathcal{N}_-$, $\mathcal{N}_+$	The set of energy combinations that corresponds to $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively. See after Eq. 4.26.
$\mathbf{n}$, $\mathbf{n}(\mathbf{c})$, $\mathbf{n}(\tilde{\mathbf{c}})$	Energy combination an $\mathbf{H}$ eigenstate with energy configuration $\mathbf{c}$. Same for $\mathbf{n}(\tilde{\mathbf{c}})$. See Def. 4.3.5.
$\mathbf{O}_{jk}^{(i)}$	The jk-th block of the perturbation $\mathbf{V}^{(i)}$ corresponding to transition from $\mathcal{P}_j^{(i)}$ to $\mathcal{P}_k^{(i)}$, see Eq. 4.12
$\mathcal{P}_i^{(j)}$	The i-th subspace of the j-th subsystem $\mathbf{H}^{(j)}$. Sometimes also written as $\mathcal{P}_i$ if context permits.
$\mathbf{P}_i^{(j)}$	Projector onto $\mathcal{P}_i^{(j)}$. Defined in Equation 4.10.
$\mathbf{P}(\mathbf{c})$	Projector onto the subspace of each subsystem as described by energy configuration $\mathbf{c}$. See Eq. 4.26.

$\mathcal{S}_{\mathbf{n}}$, $\mathcal{S}_{\mathbf{n},\mathbf{n}'}$	Set of 4-tuples stored in the node $v_{\mathbf{n}}$ or edge $e(v_{\mathbf{n}}, v_{\mathbf{n}'})$ in $G(\mathcal{V}, \mathcal{E})$ generated in Alg. 2. See Sec. 4.4.2.
$\mathbf{T}_r$	The r-th order term in the self energy expansion $\mathbf{\Sigma}_-(z)$. See Equations 4.15 and 4.17.
$\mathbf{V}^{(i)}$	Perturbation that acts on the Hilbert space $\mathcal{H}^{(i)} \otimes \mathcal{B}$. See Figure 1a of the main text and Eq. 4.11.
$\mathbf{V}$	Total perturbation $\mathbf{H}_{\mathcal{B}} + \mathbf{V}^{(1)} + ... + \mathbf{V}^{(m)}$, see Equation 4.11
z	Expansion parameter for perturbation series. See Section 4.2.1 after Equation 4.8.

Table C.2
Table of notations (Greek alphabet) that have recurring appearances in Chapter 4.

Symbol	Meaning and first appearance
Δ	The spectral gap between the ground and the first excited state of $\mathbf{H}$. See Section 4.2.1 opening.
$\boldsymbol{\lambda}$, λ_i	Basic quantity for constructing an upper bound to $\|\mathbf{T}_r\|_2$. See Definition 4.3.2.
$\boldsymbol{\Pi}_-$, $\boldsymbol{\Pi}_+$	Projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively. See text before Equation 4.13 and also Equation 4.27.
$\|\psi_{j,p}^{(i)}\rangle$	The p-th degenerate eigenvector of $\mathcal{P}_j^{(i)}$. See Equation 4.10.
ω	Basic quantity for constructing an upper bound to $\|\mathbf{T}_r\|_2$. See Definition 4.3.1.

D. EFFICIENT ALGORITHM FOR COMPUTING MONOMIAL SYMMETRIC POLYNOMIALS

We start with a property of monomial symmetric polynomials that is instrumental to our algorithm design. Although the proof is rather elementary, we state it in order to facilitate further discussions.

Lemma D.0.1 *Consider monomial symmetric polynomial* $m_{\mathbf{b}} : \mathbb{R}^n \mapsto \mathbb{R}$ *with* $\mathbf{b} \in \mathbb{N}^k$ *as its reduced partition (cf. Definition 4.2.2). Let* $\mathbf{x} = (x_1, x_2, \cdots, x_n)$. *Then for any positive integer* s

$$m_{\mathbf{b}}(\mathbf{x}) m_{(s)}(\mathbf{x}) = \begin{cases} m_{(s,\mathbf{b})}(\mathbf{x}) + \sum_{t=1}^{k} m_{\mathbf{b}'_t}(\mathbf{x}) & k < n \\ \sum_{t=1}^{k} m_{\mathbf{b}'_t}(\mathbf{x}) & k = n \end{cases} \tag{D.1}$$

where $\mathbf{b}'_t = (b_1, b_2, \cdots, b_t + s, \cdots, b_k)$ *and* $(s, \mathbf{b})$ *denotes a new* $(k+1)$*-dimensional reduced partition vector with the element 1 concatenated to the original reduced partition* $\mathbf{b}$.

Proof For every $t \in [n]$, we could always rewrite $m_{\mathbf{b}}$ as

$$\begin{aligned} m_{\mathbf{b}}(\mathbf{x}) &= \textstyle\sum_{\pi} x_{\pi(1)}^{b_1} x_{\pi(2)}^{b_2} \cdots x_{\pi(k)}^{b_k} \\ &= \textstyle\sum_{i=1}^{k} x_t^{b_i} \sum_{\pi_t^{(i)}} x_{\pi_t^{(i)}(1)}^{b_1} \cdots x_{\pi_t^{(i)}(i-1)}^{b_{i-1}} x_{\pi_t^{(i)}(i+1)}^{b_{i+1}} \cdots x_{\pi_t^{(i)}(k)}^{b_k} \\ &\equiv f(t). \end{aligned} \tag{D.2}$$

Here $\pi : [n] \mapsto [k]$ takes k distinct elements from $[n]$ and arranges them. $\pi_t^{(i)} : [n] \backslash \{t\} \mapsto [k] \backslash \{i\}$ chooses $(k-1)$ elements from $[n] \backslash \{t\}$ and permutes them. Since

the $f(t)$ defined in Equation (D.2) satisfies $f(t) = m_{\mathbf{b}}(\mathbf{x})$ for all $t \in [n]$, the product $m_{\mathbf{b}}(\mathbf{x})m_{(s)}(\mathbf{x})$ becomes

$$m_{\mathbf{b}}(\mathbf{x})m_{(s)}(\mathbf{x}) = \sum_{t=1}^{n} x_t^s f(t) = \sum_{t=1}^{n}\sum_{i=1}^{k} x_t^{b_i+s} \sum_{\pi_t^{(i)}} x_{\pi_t^{(i)}(1)}^{b_1} \cdots x_{\pi_t^{(i)}(i-1)}^{b_{i-1}} x_{\pi_t^{(i)}(i+1)}^{b_{i+1}} \cdots x_{\pi_t^{(i)}(k)}^{b_k}. \tag{D.3}$$

If $k = n$, using Equation (D.2), Equation (D.3) becomes

$$\sum_{t=1}^{k} m_{\mathbf{b}'_j}(\mathbf{x}) \tag{D.4}$$

where $\mathbf{b}'_j$ is defined in the statement of the Lemma. When $k < n$, $f(t)$ contains terms where x_t does not appear and hence the expression Equation (D.3) becomes

$$m_{(s,\mathbf{b})}(\mathbf{x}) + \sum_{t=1}^{k} m_{\mathbf{b}'_j}(\mathbf{x}), \tag{D.5}$$

where the additional term $m_{(s,\mathbf{b})}(\mathbf{x})$ accounts for terms that are linear in x_t in $x_t f(t)$.

■

Lemma D.0.1 suggests that we could compute any $m_{\mathbf{b}}(\mathbf{x})$ with $\mathbf{b} = (b_1, b_2, \cdots, b_k) \in \mathbb{N}^k$ and $\mathbf{x} \in \mathbb{N}^n$ recursively as follows:

$$m_{\mathbf{b}}(\mathbf{x}) = m_{(b_1,b_2,\cdots,b_{k-1})}m_{(b_k)} - \sum_{i=1}^{k-1} m_{(b_1,\cdots,b_i+b_k,\cdots,b_{k-1})}. \tag{D.6}$$

Here $m_{(b_k)}$ is a power sum that takes $O(n)$ time to compute. With each recursion the symmetric polynomials on the right hand side have partitions whose lengths that are shorter than $|\mathbf{b}|$ by 1. The first iteration generates $k-1$ terms, each of which generates $k-2$ terms in the next iteration etc, until we fully express $m_{\mathbf{b}}(\mathbf{x})$ in terms of power sums *i.e.* monomial symmetric polynomials with length of partition being equal to 1. The final expression of $m_{\mathbf{b}}(\mathbf{x})$ will consist of $O(k!)$ terms involving power sums. Hence the total cost of evaluating $m_{\mathbf{b}}(\mathbf{x})$ using the method inspired by Lemma D.0.1 costs $\tilde{O}(k!n)$. For constant degree polynomials this is $O(n)$ cost, as opposed to $O(n^k)$ in case one evaluates the terms in $m_{\mathbf{b}}(\mathbf{x})$ term by term.

E. AN EXAMPLE FOR ILLUSTRATING WALKS IN UNPERTURBED EIGENSPACES

Consider the setting described in Figure E.1 with $m = 2$ and $\ell = 2$. This means that there are in total 2 copies of identical unperturbed systems. Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be their respective Hilbert spaces. $\ell = 2$ means that each of the unperturbed systems are 3-level systems with energy levels $E_0^{(1)}$, $E_1^{(1)}$ and $E_2^{(1)}$ for system 1 and similarly for system 2, with the superscript '(1)' replaced with '(2)'. We assume the subspace $\mathcal{P}_1$ for both unperturbed systems is 2-fold degenerate with eigenstates $|\psi_{1,1}\rangle$ and $|\psi_{1,2}\rangle$, as shown in Figure E.1b. Under the basis of the unperturbed eigenstates with ordering $|\psi_{0,1}\rangle$, $|\psi_{1,1}\rangle$, $|\psi_{1,2}\rangle$, $|\psi_{2,1}\rangle$, the unperturbed Hamiltonian for each subsystems $H^{(1)}$ and $H^{(2)}$ can be written as

$$\mathbf{H}^{(1)} = \begin{pmatrix} E_0^{(1)} & & & \\ & E_1^{(1)} & & \\ & & E_1^{(1)} & \\ & & & E_2^{(1)} \end{pmatrix} \otimes \mathbf{I}_{\mathcal{H}_2}, \tag{E.1}$$

$$\mathbf{H}^{(2)} = \mathbf{I}_{\mathcal{H}_1} \otimes \begin{pmatrix} E_0^{(2)} & & & \\ & E_1^{(2)} & & \\ & & E_1^{(2)} & \\ & & & E_2^{(2)} \end{pmatrix} \tag{E.2}$$

where $\mathbf{I}$ is the identity operator of appropriate dimension. We assume that there is a (large) gap Δ between E_0 and E_1 of each subsystem and $E_* = \frac{E_0+E_1}{2}$ is the cutoff. Let the low energy subspace $\mathcal{L}_- = \mathcal{P}_0^{(1)} \otimes \mathcal{P}_0^{(2)}$. This is illustrated in Figure E.1a.

We let the components $\mathbf{V}^{(1)}$ and $\mathbf{V}^{(2)}$ of the perturbation $\mathbf{V} = \mathbf{V}^{(1)} + \mathbf{V}^{(2)}$ be such that

$$\begin{aligned}
\mathbf{V}^{(1)} &= \mathbf{B}^{(1)}_{11,01} \otimes (|\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{1,1}| + |\psi^{(1)}_{1,1}\rangle\langle\psi^{(1)}_{0,1}|) \\
&+ \mathbf{B}^{(1)}_{12,01} \otimes (|\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{1,2}| + |\psi^{(1)}_{1,2}\rangle\langle\psi^{(1)}_{0,1}|) \\
&+ \mathbf{B}^{(1)}_{11,12} \otimes (|\psi^{(1)}_{1,1}\rangle\langle\psi^{(1)}_{2,1}| + |\psi^{(1)}_{2,1}\rangle\langle\psi^{(1)}_{1,1}|) \\
&+ \mathbf{B}^{(1)}_{21,12} \otimes (|\psi^{(1)}_{1,2}\rangle\langle\psi^{(1)}_{2,1}| + |\psi^{(1)}_{2,1}\rangle\langle\psi^{(1)}_{1,2}|)
\end{aligned} \tag{E.3}$$

and $\mathbf{V}^{(2)}$ is the same as $\mathbf{V}^{(1)}$ but with all superscripts replaced with '(2)'. In matrix forms,

$$\mathbf{V}^{(1)} = \begin{pmatrix}
 & \mathbf{B}^{(1)}_{11,01} & \mathbf{B}^{(1)}_{12,01} & \\
\mathbf{B}^{(1)}_{11,10} & & & \mathbf{B}^{(1)}_{11,12} \\
\mathbf{B}^{(1)}_{21,10} & & & \mathbf{B}^{(1)}_{21,12} \\
 & \mathbf{B}^{(1)}_{11,21} & \mathbf{B}^{(1)}_{12,21} &
\end{pmatrix} \otimes \mathbf{I}, \tag{E.4}$$

$$\mathbf{V}^{(2)} = \mathbf{I} \otimes \begin{pmatrix}
 & \mathbf{B}^{(2)}_{11,01} & \mathbf{B}^{(2)}_{12,01} & \\
\mathbf{B}^{(2)}_{11,10} & & & \mathbf{B}^{(2)}_{11,12} \\
\mathbf{B}^{(2)}_{21,10} & & & \mathbf{B}^{(2)}_{21,12} \\
 & \mathbf{B}^{(2)}_{11,21} & \mathbf{B}^{(2)}_{12,21} &
\end{pmatrix}. \tag{E.5}$$

As shown in Figure E.1b, we can represent the component of $\mathbf{V}^{(i)}$ acting on $\mathcal{H}_i$ as a graph with the operator $\mathbf{B}^{(i)}_{mn,jk}$ as the "weight" of the edge that corresponds to the transition $|\psi^{(i)}_{j,m}\rangle\langle\psi^{(i)}_{k,n}|$. The factors λ_i in this case are

$$\begin{aligned}\lambda_1 &= \max\{\|\mathbf{B}^{(1)}_{11,01}\|_\infty, \|\mathbf{B}^{(1)}_{12,01}\|_\infty, \|\mathbf{B}^{(1)}_{11,12}\|_\infty, \|\mathbf{B}^{(1)}_{21,12}\|_\infty\},\\ \lambda_2 &= \max\{\|\mathbf{B}^{(2)}_{11,01}\|_\infty, \|\mathbf{B}^{(2)}_{12,01}\|_\infty, \|\mathbf{B}^{(2)}_{11,12}\|_\infty, \|\mathbf{B}^{(2)}_{21,12}\|_\infty\}.\end{aligned} \tag{E.6}$$

From the diagram we could see that to excite the eigenstate $|\psi_{0,1}\rangle$ of $\mathcal{P}_0$ into $\mathcal{P}_1$, there are in total 2 ways: $|\psi_{0,1}\rangle \to |\psi_{1,1}\rangle$ and $|\psi_{0,1}\rangle \to |\psi_{1,2}\rangle$. Hence $M_{01} = 2$. Following a similar line of argument we can see that $M_{10} = 1$, $M_{12} = 1$, and $M_{21} = 2$. Because we assume that $\mathbf{V}$ is block tridiagonalizable with respect to any subsystem i, there will not be any transition from $\mathcal{P}_0$ to $\mathcal{P}_2$.

The projections of $\mathbf{V}_+$ then can be determined by taking the subgraphs in Figure E.1b on the eigenstates that belong to $\mathcal{L}_+$:

$$\begin{aligned}\mathbf{V}_+ &= \mathbf{B}^{(1)}_{11,12} \otimes (|\psi^{(1)}_{1,1}\rangle\langle\psi^{(1)}_{2,1}| + |\psi^{(1)}_{2,1}\rangle\langle\psi^{(1)}_{0,1}|)\\ &+ \mathbf{B}^{(1)}_{21,12} \otimes (|\psi^{(1)}_{1,2}\rangle\langle\psi^{(1)}_{2,1}| + |\psi^{(1)}_{2,1}\rangle\langle\psi^{(1)}_{1,2}|)\\ &+ \mathbf{B}^{(2)}_{11,12} \otimes (|\psi^{(2)}_{1,1}\rangle\langle\psi^{(2)}_{2,1}| + |\psi^{(2)}_{2,1}\rangle\langle\psi^{(2)}_{0,1}|)\\ &+ \mathbf{B}^{(2)}_{21,12} \otimes (|\psi^{(2)}_{1,2}\rangle\langle\psi^{(2)}_{2,1}| + |\psi^{(2)}_{2,1}\rangle\langle\psi^{(2)}_{1,2}|).\end{aligned} \tag{E.7}$$

The projections $\mathbf{V}_{-+}$ (resp. $\mathbf{V}_{+-}$) are respectively cuts of edges that go from $\mathcal{L}_-$ to $\mathcal{L}_+$ (resp. $\mathcal{L}_+$ to $\mathcal{L}_-$):

$$\begin{aligned}\mathbf{V}_{-+} &= \mathbf{B}^{(1)}_{11,01} \otimes |\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{1,1}| + \mathbf{B}^{(1)}_{12,01} \otimes |\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{1,2}|\\ &+ \mathbf{B}^{(2)}_{11,01} \otimes |\psi^{(2)}_{0,1}\rangle\langle\psi^{(2)}_{1,1}| + \mathbf{B}^{(2)}_{12,01} \otimes |\psi^{(2)}_{0,1}\rangle\langle\psi^{(2)}_{1,2}|\\ \mathbf{V}_{+-} &= \mathbf{B}^{(1)}_{11,10} \otimes |\psi^{(1)}_{1,1}\rangle\langle\psi^{(1)}_{0,1}| + \mathbf{B}^{(1)}_{21,10} \otimes |\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{1,2}|\\ &+ \mathbf{B}^{(2)}_{11,10} \otimes |\psi^{(2)}_{0,1}\rangle\langle\psi^{(2)}_{1,1}| + \mathbf{B}^{(2)}_{21,10} \otimes |\psi^{(2)}_{0,1}\rangle\langle\psi^{(2)}_{1,2}|.\end{aligned} \tag{E.8}$$

The operator valued resolvent $\mathbf{G}_+(z) = (z\mathbf{I} - \mathbf{H})^{-1}$ could then be written as

$$\begin{aligned}
\mathbf{G}_+(z) &= \frac{1}{z-E_1}(|\psi^{(1)}_{1,1}\rangle\langle\psi^{(1)}_{1,1}| + |\psi^{(1)}_{1,2}\rangle\langle\psi^{(1)}_{1,2}|) \otimes |\psi^{(2)}_{0,1}\rangle\langle\psi^{(2)}_{0,1}| \\
&+ \frac{1}{z-E_1}|\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{0,1}| \otimes (|\psi^{(2)}_{1,1}\rangle\langle\psi^{(2)}_{1,1}| + |\psi^{(2)}_{1,2}\rangle\langle\psi^{(2)}_{1,2}|) \\
&+ \frac{1}{z-2E_1}(|\psi^{(1)}_{1,1}\rangle\langle\psi^{(1)}_{1,1}| + |\psi^{(1)}_{1,2}\rangle\langle\psi^{(1)}_{1,2}|) \otimes (|\psi^{(2)}_{1,1}\rangle\langle\psi^{(2)}_{1,1}| + |\psi^{(2)}_{1,2}\rangle\langle\psi^{(2)}_{1,2}|) \\
&+ \frac{1}{z-E_2}(|\psi^{(1)}_{2,1}\rangle\langle\psi^{(1)}_{2,1}| \otimes |\psi^{(2)}_{0,1}\rangle\langle\psi^{(2)}_{0,1}| + |\psi^{(1)}_{0,1}\rangle\langle\psi^{(1)}_{0,1}| \otimes |\psi^{(2)}_{2,1}\rangle\langle\psi^{(2)}_{2,1}|)
\end{aligned} \tag{E.9}$$

In our projector notations, we could rewrite $\mathbf{G}_+$ as

$$\begin{aligned}
\mathbf{G}_+(z) &= \frac{1}{z-E_1}(\mathbf{P}^{(1)}_1 \otimes \mathbf{P}^{(2)}_0 + \mathbf{P}^{(1)}_0 \otimes \mathbf{P}^{(2)}_1) + \frac{1}{z-2E_1}\mathbf{P}^{(1)}_1 \otimes \mathbf{P}^{(2)}_1 \\
&+ \frac{1}{z-2E_1}(\mathbf{P}^{(1)}_2 \otimes \mathbf{P}^{(2)}_0 + \mathbf{P}^{(1)}_0 \otimes \mathbf{P}^{(2)}_2) \\
&= \frac{1}{z-E_1}(\mathbf{P}([1,0]) + \mathbf{P}([0,1])) + \frac{1}{z-2E_1}\mathbf{P}([1,1]) \\
&+ \frac{1}{z-E_2}(\mathbf{P}([2,0]) + \mathbf{P}([0,2])).
\end{aligned} \tag{E.10}$$

With definitions in eqs. (E.7) to (E.9) we could express any r-th order term $\mathbf{T}_r = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{r-2}\mathbf{G}_+\mathbf{V}_{+-}$ as a sum of terms involving $\mathbf{B}^{(i)}_{mn,jk}$ operators. For example,

$$\begin{aligned}
\mathbf{T}_2 &= \mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-} = \frac{1}{z-E_1}\Big(\mathbf{B}^{(1)}_{11,01}\mathbf{B}^{(1)}_{11,10} + \mathbf{B}^{(1)}_{12,01}\mathbf{B}^{(1)}_{21,10} \\
&+ \mathbf{B}^{(2)}_{11,01}\mathbf{B}^{(2)}_{11,10} + \mathbf{B}^{(2)}_{12,01}\mathbf{B}^{(2)}_{21,10}\Big) \otimes \mathbf{\Pi}_-.
\end{aligned} \tag{E.11}$$

Note in (E.11) that there are in total four terms, two for each subsystem. The fact that there are two terms for each subsystem is due to the fact that for each subsystem there are at most two ways to transform, through perturbation V, an eigenstate (of H) in $\mathcal{P}_0$ to one in $\mathcal{P}_1$ (Figure E.1b). In other words, $M_{01} = 2$. For an eigenstate in $\mathcal{P}_1$, there are at most one way to be transformed into $\mathcal{P}_0$ or $\mathcal{P}_2$ (or in other words,

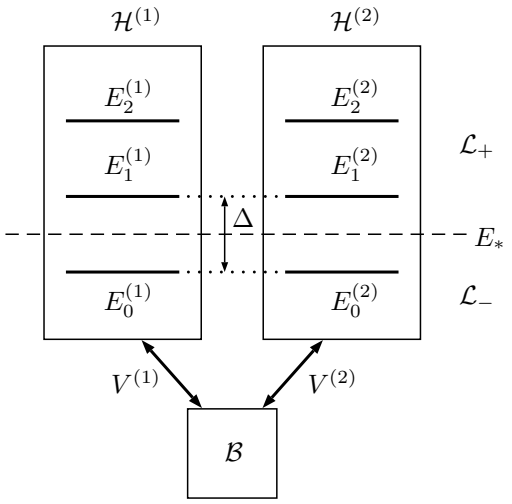
$M_{10} = 1$ and $M_{12} = 1$). Applying the definitions of λ_i, we have an upper bound to the ∞-norm of T_2 as

$$\|\mathbf{T}_2\|_\infty = \|\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-}\|_\infty \leq \frac{1}{z - E_1} 2(\lambda_1^2 + \lambda_2^2) = \frac{1}{z - E_1} M_{01} M_{10} m_{(2)}. \tag{E.12}$$

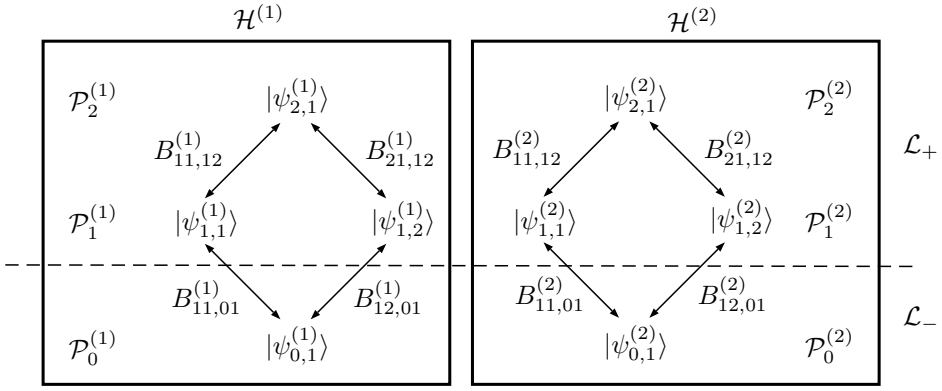
The upper bound in the above equation can be interpreted diagrammatically as in Figure E.2. The diagram shows how the upper bound to the ∞-norm "evolve" as we compute the upper bounds to $\|\mathbf{V}_{-+}\|_\infty$, $\|\mathbf{V}_{-+}\mathbf{G}_+\|_\infty$, and $\|\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_{+-}\|_\infty$:

$$\begin{aligned} \|\mathbf{V}_{-+}\|_\infty &\leq 2(\lambda_1 + \lambda_2) = M_{01} m_{(1)} \\ \|\mathbf{V}_{-+}\mathbf{G}_+\|_\infty &\leq \frac{1}{z - E_1} \cdot 2(\lambda_1 + \lambda_2) = \frac{1}{z - E_1} M_{01} m_{(1)} \end{aligned} \tag{E.13}$$

and an upper bound to $\|\mathbf{T}_2\|_\infty$ is computed in (E.12).



(a)



(b)

Fig. E.1. An example for illustrating the setting of perturbation theory that is concerned in this work.

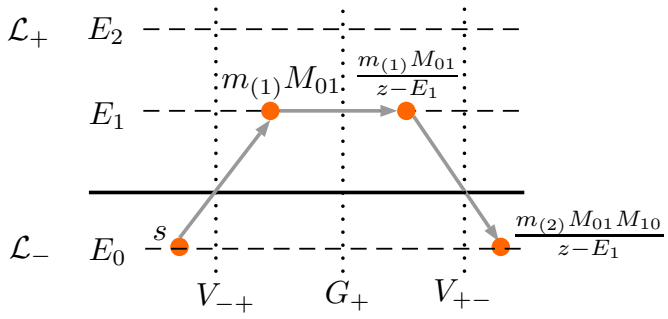


Fig. E.2. Diagram illustrating the virtual transitions associated with $\mathbf{T}_2$. Here each horizontal line represents an (unperturbed) energy level. Each vertical line represents an operator in $\mathbf{T}_r$ (here we show the diagram for $r = 2$). Each edge is associated both horizontally with an energy level and vertically with the operator corresponding to the vertical line that the edge crosses. Each node is associated with an upper bound to $\|\mathbf{Q}_{e_1}\mathbf{Q}_{e_2}\cdots\mathbf{Q}_{e_k}\|_\infty$ with $e_1, \cdots, e_k$ forming a path from the starting node s to the current node and $\mathbf{Q}_e$ being the operator associated with edge e.

F. AN EXAMPLE FOR ILLUSTRATING WALKS IN REDUCED CONFIGURATIONS

Lemma 4.3.3 has established the basic idea that $\mathbf{T}_r$ is essentially a sum of operator products associated with specific types of walks in the space of energy configuration $\mathbf{c}$. For each particular walk, we could bound the ∞-norm of its corresponding operator product using a product of scalar quantities λ_i, M_{jk} introduced in Definition 4.3.2 and 4.3.3 and $\frac{1}{z-E}$ where E is taken from the set described in Equation (4.9). For a setting with m unperturbed subsystems, $\mathbf{T}_r$ is a summation of contributions from $O(m^r)$ walks. For example in $\mathbf{T}_r = \mathbf{V}_{-+}(\mathbf{G}_+\mathbf{V}_+)^{r-2}\mathbf{G}_+\mathbf{V}_{+-}$ for any r, the first factor $\mathbf{V}_{-+}$ corresponds to the first step in the walk that departs from $\mathcal{L}_-$ into $\mathcal{L}_+$. To accomplish such departure one could excite any of the m subsystems to raise the total energy into the high energy subspace $\mathcal{L}_+$, which gives a sum

$$\lambda_1 M_{01} + \cdots + \lambda_m M_{01} \tag{F.1}$$

as shown in Equation (E.13). Each term in the sum corresponds to a distinct walk. If we consider the lowest order term T_2, which sums over contributions from 2-step walks that first enters $\mathcal{L}_+$ and immediately return to $\mathcal{L}_-$, each walk that contributes to T_2 must first excite a subsystem and subsequently de-excite it so that the total state returns to $\mathcal{L}_-$. Hence an upper bound to $\|\mathbf{T}_2\|_\infty$ can be computed as

$$\frac{1}{z-E_1}[(\lambda_1 M_{01})(\lambda_1 M_{10}) + (\lambda_2 M_{01})(\lambda_2 M_{10}) + \cdots + (\lambda_m M_{01})(\lambda_m M_{10})] \tag{F.2}$$

where E_1 is the first energy level above the cutoff λ_*. Expression F.2 is identical to the right hand side of Equation E.12 in the Appendix E, where a far more detailed derivation is presented. Expression F.2 is written in a way that highlights the structure of a summation over contributions from 2-step walks. The term in each pair of parenthesis $(\cdot)$ corresponds to the factor contributed from a single step. For general

$\mathbf{T}_r$ we have $O(r^m)$ products of such $(\cdot)$ terms to sum over, which could quickly become computationally infeasible for large systems. Using symmetric polynomials to represent the summation, as can be seen in Equations E.12 and E.13, alleviates this concern by turning the problem of managing expressions such as Equations F.1 and F.2 into the problem managing the reduced partitions (Definition 4.2.2) of symmetric polynomials. The process of summing over walks in $\mathbf{c}$ hence becomes summing over walks in the space of reduced configurations $\tilde{\mathbf{c}}$.

We now consider 4-th order perturbation theory *i.e.* $r = 4$. Figure F.1 illustrates the process of finding an upper bound to $\|\mathbf{T}_4\|_\infty$ according to Lemma 4.3.6. There are in total 3 distinct walks in $\tilde{\mathbf{c}}$ and indeed the upper bound of $\|\mathbf{T}_r\|_\infty$, denoted as $\triangle$ in Figure F.1, consists of 3 terms of symmetric polynomials with distinct partitions. Each step of the walk is driven by an operator in $\mathbf{T}_r$. Each node that the walk passes through corresponds to both a specific energy configuration and a particular position in the walk. Each node is also associated with a scalar number that serves as an upper bound to the ∞-norm of the product of operators so far.

An analogous diagram for $\mathbf{T}_2$ is shown in Figure E.2 in Appendix E. The upper bounds associated with the nodes passed through by the walk undergo a certain kind of "evolution" as the walk progresses, as can be observed both Figures E.2 and F.1. Informally the "evolution" can be described as the following: we start from an upper bound for $\|\mathbf{V}_{-+}\|_\infty$. By modifying the upper bound according to some fixed rules, we arrive at an upper bound for $\|\mathbf{V}_{-+}\mathbf{G}_+\|_\infty$. Then by further modifying the upper bound for $\|\mathbf{V}_{-+}\mathbf{G}_+\|_\infty$ we get an upper bound for $\|\mathbf{V}_{-+}\mathbf{G}_+\mathbf{V}_+\|_\infty$ etc.

The goal of the algorithms presented in the Sections 4.4.1 and 4.4.2 is to efficiently automate this "evolution" of walks using cellular automaton as the basic data structure. In the context of Algorithm 4, each horizontal line in Figure F.1 corresponds to a cell (or a node) of the graph $G(\mathcal{V}, \mathcal{E})$ generated by BuildCA in Algorithm 2 and each vertical column of nodes corresponds to a snapshot of the cell states at a given repetition of cell updates during step 2i of PerturbBound in Algorithm 4. An upper bound for $\|\mathbf{T}_r\|_\infty$ is computed by evolving the cellular automaton r times

in total $(r-1)$ times during step 2i and once during step 2k). Each path in Figure F.1 corresponds to a walks in $\tilde{\mathbf{c}}$, which by Theorem 4.4.1, also corresponds to a trace of the algorithm.

Another observation concerns the property of monomial symmetric polynomials. Note first that $m_{\mathbf{b}}(\boldsymbol{\lambda})$ contains terms that have one-one correspondence with walks that consists of b_1 transitions on one subsystem, b_2 transitions on another system, b_3 transitions on another system etc. For example, if we have $m=3$ subsystems, then $\boldsymbol{\lambda}=(\lambda_1,\lambda_2,\lambda_3)$ and the symmetric polynomial $m_{(1,3)}(\boldsymbol{\lambda})=\lambda_1\lambda_2^3+\lambda_1\lambda_3^3+\lambda_2\lambda_1^3+\lambda_2\lambda_3^3+\lambda_3\lambda_1^3+\lambda_3\lambda_2^3$ represents a collection of 4-step walks (because the sum of elements in the reduced partition is 4). Each term in $m_{(1,3)}(\boldsymbol{\lambda})$ corresponds to a type of 4-step walk. If we consider 5-step walks that are continuation of 4-step walks included in $m_{(1,3)}(\boldsymbol{\lambda})$, naturally we could choose any subsystem to act on for the 5-th step. An algebraic way of describing this freedom of choice is to use the sum $\lambda_1+\lambda_2+\lambda_3=m_{(1)}(\boldsymbol{\lambda})$. Hence the collection of 5-step walks with the first 4 steps being any walk contained $m_{(1,3)}(\boldsymbol{\lambda})$ can be represented as [178, Lemma 1]

$$m_{(1,3)}(\boldsymbol{\lambda})m_{(1)}(\boldsymbol{\lambda})=m_{(2,3)}(\boldsymbol{\lambda})+m_{(1,4)}(\boldsymbol{\lambda})+m_{(1,1,3)}(\boldsymbol{\lambda}). \tag{F.3}$$

The above equation shows an example of generating terms for $(t+1)$-step walks from terms for t-step walks. As can be noticed from Figure F.1, such "generation" mechanism of high-order symmetric polynomials from lower-order ones as exemplified in Equation F.3 plays an important role in the "evolution" of upper bounds mentioned in the previous paragraph.

If one runs PerturbBound$(4,\boldsymbol{\lambda},\mathbf{M})$ as described in Algorithm 4 with the initial assignment of cell state being $\mathbf{n}_-=\mathbf{n}_0$ and $\mathbf{n}_+=(m-1,1,0,\cdots,0)$ during step 2a through 2g, the returned value $\tau_{4,\mathbf{n}_-}$ at step 2l should be the total value of the list of 4-tuples shown in Table F.1, which is

$$\begin{aligned}\tau_{4,\mathbf{n}_-}=\sum_{\mathcal{T}=(\tilde{\mathbf{c}},\mathbf{b},\xi,\mu)\in\mathcal{S}_{\mathbf{n}_0}} m_{\mathbf{b}}(\boldsymbol{\lambda}) &= \frac{M_{01}M_{10}\omega^2}{(z-E_1)^3}m_{(2)}+\frac{2M_{01}^2M_{10}^2}{(z-E_1)^2(z-2E_1)}m_{(2,2)}\\ &\quad+\frac{M_{01}M_{12}M_{21}M_{10}}{(z-E_1)^2(z-E_2)}m_{(4)}=(\triangle)\end{aligned} \tag{F.4}$$

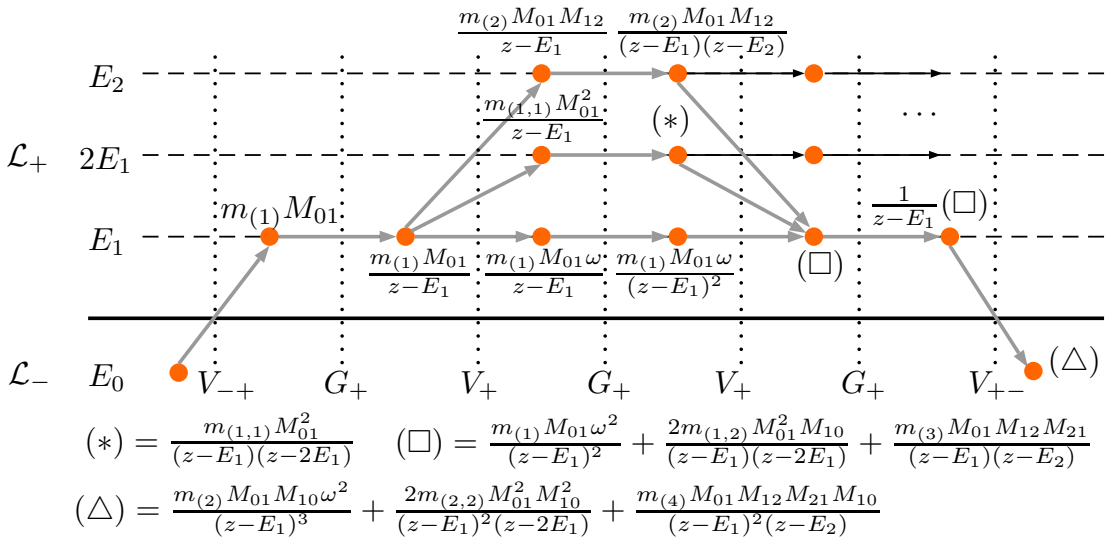


Fig. F.1. An example of enumerating 4-step walks in $\tilde{\mathbf{c}}$. Each path marked with bold edges corresponds to a walk in $\tilde{\mathbf{c}}$ with $\tilde{\mathbf{c}}^{(0)} = \tilde{\mathbf{c}}^{(4)} = (0, \cdots, 0)$. Due to limited space we replace some of the longer expressions with symbols $(*)$, $(\square)$ and $(\triangle)$ in the diagram and provide their full expressions below the diagram. Here we assume that $\mathcal{L}_-^{(i)} = \mathcal{P}_0^{(i)}$ for any i. Each horizontal line represents an energy level of the total unperturbed system $\mathcal{H}^{(1)} \otimes \mathcal{H}^{(2)} \otimes \cdots \otimes \mathcal{H}^{(m)}$, or equivalently an energy combination $\mathbf{n}$. Each vertical line represents an operator in $\mathbf{T}_r$. Each edge is associated both horizontally with an energy level and vertically with the operator corresponding to the vertical line that the edge crosses.

where by $(\triangle)$ we refer to Figure F.1. Equation F.4 is also one of the terms on the right hand side of Equation 4.43 in Lemma 4.3.6 with $\tilde{\mathbf{c}}^{(0)} = \tilde{\mathbf{c}}^{(4)} = (0, \cdots, 0)$. See Figure F.1.

Table F.1
4-tuple list associated with the cell $\mathcal{S}_{\mathbf{n}_0}$, representing the the expression ($\triangle$) in Figure F.1, which is the final upper bound computed for $\|\mathbf{T}_4\|_\infty$.

$\tilde{\mathbf{c}}$	$\mathbf{b}$	ξ	μ
(0)	(2)	$\frac{M_{01}M_{10}\omega^2}{(z-E_1)^3}$	$\tilde{\mathbf{c}} = (0)$ $\downarrow$ $\mathbf{b} = (2)$
(0,0)	(2,2)	$\frac{2M_{01}^2M_{10}^2}{(z-E_1)^2(z-2E_1)}$	$\tilde{\mathbf{c}} = (0 \quad 0)$ $\downarrow \quad \downarrow$ $\mathbf{b} = (2 \quad 2)$
(0)	(4)	$\frac{M_{01}M_{12}M_{21}M_{10}}{(z-E_1)^2(z-E_2)}$	$\tilde{\mathbf{c}} = (0)$ $\downarrow$ $\mathbf{b} = (4)$

VITA

VITA

Yudong Cao obtained his Bachelor of Science in mechanical engineering in spring 2011, simultaneously from Purdue University and Shanghai Jiaotong University due to an exchange program between the two institutions. In spring 2013, he finished his master's degree, also in mechanical engineering at Purdue, with a thesis on quantum algorithm for solving linear systems of equations. He joined the Computer Science Department in Fall 2013. In spring 2015, obtained a second master's degree, which is in computer science. He defended his doctoral dissertation in August 2016. Immediately after completing the program, he started working at Harvard University as a postdoctoral researcher.