# Towards Viable Quantum Computation for Chemistry

# Towards Viable Quantum Computation for Chemistry

A dissertation presented
by
Ryan Joseph Babbush
to
The Committee in Chemical Physics

in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy
in the subject of
Chemical Physics

Harvard University
Cambridge, Massachusetts

May 2015

Dissertation Advisor

**Alán Aspuru-Guzik**

Author

**Ryan Joseph Babbush**

**Towards Viable Quantum Computation for Chemistry**

## Abstract

Since its introduction one decade ago, the quantum algorithm for chemistry has been among the most anticipated applications of quantum computers. However, as the age of industrial quantum technology dawns, so has the realization that even "polynomial" resource overheads are often prohibitive. There remains a large gap between the capabilities of existing hardware and the resources required to quantum compute classically intractable problems in chemistry. The primary contribution of this dissertation is to take meaningful steps towards reducing the costs of three approaches to quantum computing chemistry. First, we discuss how chemistry problems can be embedded in Hamiltonians suitable for commercially manufactured quantum annealing machines. We introduce schemes for more efficiently compiling problems to annealing Hamiltonians and apply the techniques to problems in protein folding, gene expression, and cheminformatics. Second, we introduce the first adiabatic quantum algorithm for fermionic simulation. Towards this end, we develop tools which embed arbitrary universal Hamiltonians in constrained hardware at a reduced cost. Finally, we turn our attention to the digital quantum algorithm for chemistry. By exploiting the locality of physical interactions, we quadratically reduce the number of terms which must be simulated. By analyzing the scaling of time discretization errors in terms of chemical properties, we obtain significantly tighter bounds on the minimum number of time steps which must be simulated. Also included in this dissertation is a protocol for preparing configuration interaction states that is asymptotically superior to all prior results and the details of the most accurate experimental quantum simulation of chemistry ever performed.

# Contents

## Citations to previously published work

Apart from minor modifications, Chapters 2-10 appeared as the following publications:

"Resource Efficient Gadgets for Compiling Adiabatic Quantum Optimization Problems". Ryan Babbush, Bryan O'Gorman, and Alán Aspuru-Guzik. *Annalen der Physik*. Volume 525, Number 10-11: 877-888. 2013.

"Construction of Energy Functions for Lattice Heteropolymer Models: Efficient Encodings for Constraint Satisfaction Programming and Quantum Annealing". Ryan Babbush, Alejandro Perdomo-Ortiz, Bryan O'Gorman and Alán Aspuru-Guzik. *Advances in Chemical Physics*. Volume 155, Chapter 5: 201-243. 2014.

"Bayesian Network Structure Learning Using Quantum Annealing". Bryan O'Gorman, Alejandro Perdomo-Ortiz, Ryan Babbush, Alán Aspuru-Guzik and Vadim Smelyanskiy. *European Physical Journal Special Topics*. Volume 225, Number 1: 163-188. 2015.

"Construction of Non-Convex Polynomial Loss Functions for Training a Binary Classifier with Quantum Annealing". Ryan Babbush, Vasil Denchev, Nan Ding, Sergei Isakov and Hartmut Neven. arXiv preprint 1406.4203. 1-9. 2014.

"Hamiltonian Gadgets with Reduced Resource Requirements". Yudong Cao, Ryan Babbush, Jacob Biamonte, and Sabre Kais. *Physical Review A*. Volume 91, Number 1: 012315. 2015.

"Adiabatic Quantum Simulation of Quantum Chemistry". Ryan Babbush, Peter Love and Alán Aspuru-Guzik. *Scientific Reports*. Volume 4, Number 6603: 1-11. 2014.

"Chemical Basis of Trotter-Suzuki Errors in Quantum Chemistry Simulation". Ryan Babbush, Jarrod McClean, Dave Wecker, Alán Aspuru-Guzik and Nathan Wiebe. *Physical Review A*. Volume 91, Number 2: 022311. 2015.

"Exploiting Locality in Quantum Computation for Quantum Chemistry". Jarrod McClean, Ryan Babbush, Peter Love and Alán Aspuru-Guzik. *Journal of Physical Chemistry Letters*. Volume 5, Number 24: 4368-4380. 2014.

"Quantum Simulation of Helium Hydride in a Solid-State Spin Register". Ya Wang, Florian Dolde, Jacob Biamonte, Ryan Babbush, Ville Bergholm, Sen Yang, Ingmar Jakobi, Philipp Neumann, Alán Aspuru-Guzik, James D. Whitfield and Jørg Wrachtrup. arXiv preprint 1405.2696. 1-9. 2014.

I contributed to two additional publications during the course of my doctorate which are not related to quantum computation and thus did not make it into this dissertation [17, 91]:

> "Force-Field Functor Theory: Classical Force-Fields which Reproduce Equilibrium Quantum Distributions". Ryan Babbush, John Parkhill, and Alán Aspuru-Guzik. *Frontiers in Chemistry.* Volume 1, Number 26: 1-10. 2013.

> "Bayesian Sampling Using Stochastic Gradient Thermostats". Nan Ding, Youhan Fang, Ryan Babbush, Changyou Chen, Robert Skeel and Hartmut Neven. *Advances in Neural Information Processing Systems.* Number 27: 3203-3211. 2014.

## Acknowledgments

First and foremost, I would like to thank my advisor, teacher, and friend, Professor Alán Aspuru-Guzik. Alán's infectious enthusiasm for science, his endless excitement and his unpretentious demeanor foster an environment where students learn because there are no stupid questions and where creativity abounds because discussing science is as much fun as it is serious. Working in Alán's lab provided me the opportunity to pursue the projects I found most interesting, work with the people from whom I could learn the most, and seek out opportunities (such as internships at Microsoft and Google) which advanced a non-academic scientific career. Though I've learned a lot of science from Alán, by watching him I think I have learned even more about how to inspire people and be a good boss.

Next, I thank three of my senior coauthors whose sustained mentorship has had very significant impact on my development as a scientist. During my second year, I worked closely with Professor John Parkhill (then a postdoc) who taught me much about programming, statistical mechanics, and research methods in general. While interning at Microsoft Research, I spent hours each day pestering Dr. Nathan Wiebe who responded by teaching me almost everything I know about universal quantum simulation. Last but not least, it has been a true privilege to work closely with Professor Peter Love for the last two years. Peter's constant advice on everything from the politics of industry to the classification of Lie algebras has contributed substantially to my success in graduate school.

I am also grateful for my affiliation with the Quantum A.I. Lab, which is joint between Google and NASA. Throughout my internship and collaborations with the group at Google I learned much about quantum annealing from my many conversations with Dr. Sergio Boixo, Dr. Masoud Mohseni, Dr. Alireza Shabani, Dr. Sergei Isakov, Dr. Hartmut Neven

and Professor Eddie Farhi. I similarly learned much from the patient machine learning specialists in the Google group, Dr. Nan Ding and Dr. Vasil Denchev, who are also responsible for making sure I knew enough about algorithms to pass the engineering interviews at Google. I also had the pleasure of working extensively with two members of the NASA group: Dr. Alejandro Perdomo-Ortiz who was at Harvard during my first year, and Bryan O'Gorman who was at Harvard during my second year.

The summer after my second year I had the opportunity to visit and collaborate with Dr. Jacob Biamonte at the Institute for Scientific Interchange in Torino, Italy. Among other lessons, Jacob taught me much about perturbation theory and introduced me to Yudong Cao. Subsequently, Jacob and Yudong invited me to help them complete a paper on perturbative gadgets. During my last winter in Boston, I was invited to visit Professor Rainer Blatt's group at the Institute for Quantum Optics and Quantum Information in Innsbruck, Austria. There, I had the rewarding opportunity to work on an experimental collaboration with Dr. Cornelius Hempel.

Of all my fellow students in the Aspuru-Guzik group, I would especially like to thank Jarrod McClean for collaborations on multiple papers and for patiently answering many questions with the intuition and knowledge of an experienced scientist. Other current or former Aspuru-Guzik group members who I have significantly learned from include Professor Man-Hong Yung, Professor Joel Yuen, Professor Johannes Hachmann, Professor Felipe Herrera, Dr. Semion Saikin, Dr. Dmitrij Rapport, Dr. Dmitry Zubarev, Dr. Salvatore Mandrà, Dr. Borja Peropadre, Dr. Sarah Mostame, Dr. Gian Giacomo Guerreschi, Dr. David Tempel, Dr. Xavier Andrade, Dr. Roberto Olivares, Dr. James Whitfield, Thomas Markovich, Sam Blau, Stéphanie Valleau, Nico Sawaya, Jon Welch, Jacob Sanders, Joey Goodknight, Jhonathan Romero Fontalvo, Ian Kivlichan and Annie Wei.

Prior to graduate school, there were several important scientific role models who significantly shaped my decision to end up studying quantum information. Among these, none is more important than my high school chemistry teacher and advisor, Lewis Acampora, who is largely responsible for convincing me to think seriously about a career in science. As an undergraduate at Carleton, I also benefited from excellent teachers such as Professor Daniela Kohen and Professor Arjendu Pattanayak. Other important role models include Professor Eric Heller who provided me with excellent guidance throughout my first and second years at Harvard and afterwards as a member of my dissertation committee. I am also grateful to Professor Eugene Shakhnovich for taking the time to review this dissertation and provide guidance as a member of my dissertation committee.

Last but not least, I am forever grateful to my friends and family who have and continue to support my career in science. In particular, my sister Amy Babbush, my grandmothers Judy Moore and Elaine Babbush and my grandfathers Thomas Moore and Harvey Babbush. Of all these people, none have been better role models, or more supportive and loving than my mother Suzanne Babbush and my father Randall Babbush.

*To my parents*

## Chapter 1

# Introduction

During the several hundred years leading up to the discovery of quantum mechanics it was common to think of the universe as a vast, intricate clockwork mechanism, deterministically ticking along with its gears governed by the laws of classical physics [138]. Eventually, computers replaced machines in the popular metaphors and we began to understand dynamics as a special case of information processing. Since the instantaneous state of a physical system is literally "information", any time evolution altering the system performs a computation on information encoded in the state. Accordingly, the task of modeling a physical system is equivalent to simulating a computation performed by nature. We might wonder then, how powerful a computer is nature? Ignoring a few technicalities having to do with analog computation and chaos, the answer is that nature seems to process information just like any other computer, except when things get really small or really cold, that is, when things get quantum. This is because merely representing the state of a quantum system on a classical computer appears to require an amount of information that increases exponentially with the size of the system. Such systems are said to store "quantum information".

For more than fifty years after the discovery of quantum mechanics, the exponential classical overhead required to simulate quantum systems was, to some, a source of despair.

Paul Dirac perfectly summed up the situation in 1929 [92],

> The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

It did not occur to Dirac that instead of surrendering to the complexity of quantum mechanics, we might harness it as a computational resource. In 1982, Richard Feynman first[1] suggested that we could build a computer which operated according to the laws of quantum mechanics in order to efficiently simulate quantum dynamics [107]. The field of quantum computing would remain rather obscure until 1994 when Peter Shor revealed a quantum algorithm for integer factorization (a problem having nothing to do with physics) which is exponentially more efficient than the best known classical algorithm [251].

Several papers in the late nineties would formalize Feynman's notion of a universal quantum computer capable of simulating arbitrary quantum dynamics. In 1995, Alexei Kitaev published a general method (today known as the quantum phase estimation algorithm) for efficiently estimating the eigenvalues of unitary operators such as those which evolve quantum systems forward in time [168]. The following year, Seth Lloyd showed that any local quantum system could be efficiently simulated using a set of standard quantum logic gates [183]. Subsequent work by Abrams and Lloyd demonstrated how this could be done for fermions [1] and provided a method for combining phase estimation and quantum simulation in order to find the eigenvalues and eigenvectors of a quantum system [2]. While providing a prescription for evolving quantum states under arbitrary Hamiltonians, these papers did not demonstrate how one might efficiently prepare physical states of interest in the first place.

[1]Though unknown to most Western scientists at the time, Yuri Manin made a similar suggestion in paper published in Russian two years prior [189].

In 2000, Farhi *et al.* introduced a novel paradigm for quantum computation, known as adiabatic quantum computing [104]. In this model, a quantum system is initialized in the ground state of a trivial Hamiltonian and then the Hamiltonian is slowly perturbed into a Hamiltonian whose ground state encodes the solution to a nontrivial computational problem. If the perturbation acts sufficiently slowly, the system will remain in its instantaneous ground state throughout the evolution. In 2005, Aspuru-Guzik *et al.* conjectured that because the natural processes that lead to the formation of molecules are generally efficient, adiabatic state preparation can be used to efficiently prepare the ground states of chemical Hamiltonians [11]. With this assumption, Aspuru-Guzik *et al.* demonstrated that by combining adiabatic state preparation, Kitaev's quantum phase estimation algorithm and Lloyd's scheme for universal simulation, one can efficiently compute the ground state energy of molecules, thus solving the central problem of quantum chemistry known as the electronic structure problem [11].

Ten years have passed since Aspuru-Guzik *et al.*'s seminal work and quantum computers now appear to be in the very early stages of industrialization. Today there is optimism that we are technologically close to the ability to manufacture a coherent quantum annealing machine or universal quantum computer with hundreds of high quality qubits. This dissertation explores strategies for using such hypothetical "early industrial" quantum devices to have an impact on problems in chemistry. The subsequent chapters seek to realistically analyze and practically improve upon the prospects for implementing classically intractable chemistry problems on a near future quantum device. Early work in this field, such as [11], often described algorithms intended for a very abstract device with the goal of demonstrating a quantum algorithm to be exponentially faster than the classical alternative. By contrast, this dissertation is often concerned with precise sub-asymptotic scalings of resources.

## 1.1 Quantum Annealing

In the first section, this dissertation explores resource efficient strategies for encoding problems of relevance to chemistry in Hamiltonians suitable for quantum annealing. Quantum annealing is the quantum analog of simulated annealing [167, 7, 8]. In particular, it is an analog quantum algorithm which heuristically optimizes classical energy landscapes by means of quantum tunneling [108, 159, 57, 238]. In the last few years, quantum annealing has been a central focus of the quantum computing field due, in part, to the prominence of a Canadian company, D-Wave Systems, which manufactures quantum annealing machines with over 500 superconducting qubits [60]. While the extent to which the current (fairly noisy) generation of devices demonstrate a quantum advantage (if any) is the subject of considerable debate [154, 43, 225, 87, 278, 254, 174, 233, 45], their construction represents a significant step forward towards a coherent quantum annealer.

The typical implementation of quantum annealing, and the one implemented by D-Wave Systems, involves a transverse Ising Hamiltonian of the form

$$H = -A\left(t\right)\sum_{i}\sigma_{i}^{x} + B\left(t\right)H_{\text{prob}} \tag{1.1}$$

where the diagonal "problem Hamiltonian" is a programmable classical Ising model,

$$H_{\text{prob}} = -\sum_{\{i,j\}\in G} J_{ij}\sigma_{i}^{z}\sigma_{j}^{z} - \sum_{i} h_{i}\sigma_{i}^{z}. \tag{1.2}$$

Typically, the time-dependent parameters $A\left(t\right)$ and $B\left(t\right)$ are chosen so that $B\left(0\right) = A\left(T\right) = 0$ where $T$ is the total evolution time. For instance, a typical annealing schedule is to choose $A\left(t\right) = t/T$ and $B\left(t\right) = 1 - t/T$. The connectivity graph of the hardware, $G$, defines which qubits are coupled together. Engineering considerations typically limit this graph to bounded degree. For instance, D-Wave devices use the so-called "Chimera graph"

which is approximately a 5-regular graph [273, 164]. One programs a quantum annealer via specification of the local fields $h_i$ and couplings $J_{ij}$ defined over the graph $G$.

The system begins in the lowest eigenstate of the transverse field Hamiltonian, which is the uniform superposition state,

$$\frac{1}{\sqrt{2^N}} \bigotimes_{i=1}^{N} \left( |0\rangle_i + |1\rangle_i \right). \tag{1.3}$$

Thus, the system starts maximally delocalized over the energy landscape defined by the diagonal Hamiltonian, $H_{\text{prob}}$. As the transverse field is slowly turned off and the problem Hamiltonian is slowly turned on, the wavefunction localizes in the computational basis. The total time $T$ required to ensure the system ends up in the lowest energy eigenstate is inversely proportional to a polynomial power of the gap between the lowest eigenstate and first excited state [67]. In some cases, this gap shrinks exponentially fast with the size of the problem. Thus, in general, we do not expect that quantum annealing can efficiently solve NP-Hard optimization problems in the worst cases. However, there are known problems for which quantum annealing still affords an exponential speedup over any known classical algorithm [256]. Furthermore, there is evidence that multiqubit cotunneling is a generic computational resource which might provide an average case scaling advantage in terms of problem size or residual energy (i.e. solution quality) [201, 202, 45, 6, 212]. Even if quantum annealing is only capable of reducing the exponential prefactor in the exact scaling of the hardest instances of certain problems, this would be of significant industrial value and has obvious applications in many fields of research.

For many graphs $G$, determining the lowest energy eigenstate of the Hamiltonian in Eq. 1.2 is formally NP-Hard [23]. Accordingly, there exists a constructive procedure for efficiently embedding any optimization problem in the class NP in those models. Thus, there is nothing novel or surprising about the ability to efficiently cast a classical optimization

problem as the ground state of a quantum annealing Hamiltonian. However, in this context "efficient" has a fairly esoteric definition. NP-Hardness only guarantees that a problem can be embedded with "polynomial overhead". In practice, this overhead is often enormous and naive embedding schemes frequently require a quantum annealer with hundreds of thousands of qubits (or more) to solve a problem whose solution space can be described with hundreds of bits. Section I of this dissertation is about strategies for intelligently encoding optimization problems as the ground state of the Ising model so that classically intractable problems of relevance to chemistry can be attempted on existing quantum hardware.

Qubits are not the only resource that limits one's ability to encode optimization problems on existing hardware. As mentioned earlier, the connectivity of the qubits is determined by engineering considerations and it is unrealistic to expect that one will be able to realize a fully-connected Ising model. While it is always possible to find a minor embedding of a graph having arbitrary connectivity in any NP-Hard graph, doing so often requires expensive classical preprocessing and ultimately comes at the cost of many quantum bits [73, 75, 273]. Furthermore, engineering considerations also constrain the range of available couplings (this is known as control precision). Accordingly, Chapter 2 is about practical strategies for compiling arbitrary optimization problems involving variables connected through $k$-body interactions into a form suitable for quantum annealing with minimal control precision overhead.

In Chapter 3, we demonstrate several resource-efficient strategies for embedding lattice protein folding in a Hamilonian suitable for quantum annealing. This chapter significantly improves the encodings for this problem described in [220] and explains in detail the strategies used for an experimental implementation of this problem in [221]. Chapter 4 is about learning the structure of Bayesian networks, a problem of importance to understanding

metabolic pathways and gene expression networks [93, 111]. This problem represents a major bottleneck in computational genomics and the embedding strategy we employ has the desirable (and unusual) quality that the connectivity graph of the qubits is determined by problem size and does not depend on the particular instance. Chapter 5 is about a supervised machine learning problem, binary classification robust to noise, which is widely applicable in high-throughput materials or pharmaceutical screening and cheminformatics [293, 180]. This work improves upon embeddings first introduced in [85].

## 1.2 Adiabatic Quantum Computing

Adiabatic quantum computing is a variant on quantum annealing in which the system remains in the ground state throughout the entire evolution and the final Hamiltonian does not need to be diagonal. In fact, a universal adiabatic quantum computer (one capable of efficiently executing any gate model algorithm) requires the ability to implement a final Hamiltonian for which determining the ground state energy is complete for quantum computing, i.e. QMA-Complete [165, 3]. A QMA-Complete Hamiltonian is one that can embed the eigenspectrum of any Hamiltonian in the class QMA as its lowest eigenvalues while requiring (at most) polynomially more resources [165]. Ever since [165] we have known that the eigenspectrum of any class of spin Hamiltonians having fixed locality can be embedded in the low energy sector of a 2-local Hamiltonian. More recently we have known that many 2-local Hamiltonians with restricted connectivity and limited interactions are QMA-Complete [38, 216, 82]. Perhaps the simplest example of a QMA-Complete Hamiltonian is the ZZXX Ising model [38],

$$H = -\sum_{ij} J_{ij}\sigma_i^z\sigma_j^z - \sum_{ij} K_{ij}\sigma_i^x\sigma_j^x - \sum_i h_i\sigma_i^z - \sum_i g_i\sigma_i^x. \tag{1.4}$$

The electronic structure Hamiltonian can be written in second quantization as,

$$H = \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_s a_r + \sum_{pq} h_{pq} a_p^\dagger a_q \tag{1.5}$$

in which creation and annihilation operators act on a basis of orthogonal spin orbitals, $\{\varphi_i\}$ and the one-electron and two-electron integrals are

$$h_{pq} = \int d\sigma \; \varphi_p^*(\sigma) \left( -\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \varphi_q(\sigma) \tag{1.6}$$

$$h_{pqrs} = \int d\sigma_1 \; d\sigma_2 \; \frac{\varphi_p^*(\sigma_1)\varphi_q^*(\sigma_2)\varphi_s(\sigma_1)\varphi_r(\sigma_2)}{|r_1 - r_2|} \tag{1.7}$$

where $\sigma_i$ contains spatial and spin degrees of freedom for the electrons [15]. The operators $a_p^\dagger$ and $a_r$ obey the fermionic anti-commutation relations

$$\{a_p^\dagger, a_r\} = \delta_{p,r}, \qquad \{a_p^\dagger, a_r^\dagger\} = \{a_p, a_r\} = 0. \tag{1.8}$$

The primary accomplishment of Chapter 7 is to map the Hamiltonian in Eq. 1.5 to an experimentally realizable universal Hamiltonian (e.g. Eq. 1.4). This task is complicated by the fact that the typical mapping between fermions and spins, the Jordan-Wigner transformation, produces a spin Hamiltonian that has locality equal to the number of orbitals, which is to say that it produces an $N$-local Hamiltonian rather than a $k$-local Hamiltonian.

The subject of Chapter 6 is perturbative gadgets, which are tools used to embed the eigenspectra of arbitrary Hamiltonians in the low energy sector of a larger Hamiltonian that is constrained in some way [165]. For instance, perturbative gadgets can be used to embed the eigenspectra of an arbitrary $k$-local Hamiltonian in a 2-local Hamiltonian at the cost of a polynomial number of ancilla qubits. These "gadgets" were first introduced in [165] and have little in common with the classical gadgets of Chapter 2 which only work for reducing the locality of diagonal Hamiltonians but have the attractive property of embedding the energy landscape *exactly* rather than perturbatively.

Specifically, the goal of Chapter 6 is to develop improved perturbative gadgets, based on a formulation by Olivera and Terhal [216] which require significantly less experimental resources. In particular, our improved gadgets require an amount of control precision which scales asymptotically better than any other gadget construction in the literature. We also introduce perturbative gadgets which can rid a Hamiltonian of YY couplings. This result has specific utility for quantum chemistry as such terms appear in the chemistry Hamiltonian but are difficult to implement using superconducting flux quits. Due to the availability of medium-scale quantum annealers, there is optimism that a universal adiabatic quantum computer can be built using a similar superconducting flux qubit platform. To accomplish this, one needs only to add XX couplings to existing designs.

As discussed in Chapter 6, with all known formulations of perturbative gadgets, the control precision cost of embedding a $k$-local Hamiltonian in a 2-local Hamiltonian scales exponentially with $k$. Thus, to the best of our knowledge[2], if the Jordan-Wigner transformation is applied to the Hamiltonian in Eq. 1.5, the resultant $N$-local spin Hamiltonian can be reduced to a 2-local Hamiltonian only at the cost of control precision that scales exponentially in $N$. Perhaps the most important insight of Chapter 7 is that by using a technique recently described by Seeley and Love, the Bravyi-Kitaev transformation, one can map fermionic operators to spin couplings with locality that is logarithmic in $N$. Accordingly, the control precision cost of making such Hamiltonians 2-local comes at only a polynomial overhead. This insight allows us map the electronic structure Hamiltonian to an experimentally realizable qubit Hamiltonian so that we can apply the adiabatic algorithm to quantum chemistry. More generally, Chapter 7 shows the first efficient adiabatic quantum algorithm for fermionic simulation.

[2]While no such technique exists in the literature, we do not have any particular reason for suspecting that it is impossible to reduce locality with control precision overhead that scales polynomially with $k$.

## 1.3 Digital Quantum Computation

The final section of this dissertation explores techniques for solving the molecular electronic structure problem with a digital quantum computer. The canonical quantum algorithm for quantum computing molecular energies was introduced by Aspuru-Guzik *et al.* almost a decade ago [11]. Preliminary details of how one might actually implement this algorithm were explained five years later by Whitfield *et al.* [285]. Very recently, there has been substantial renewed interest in this algorithm as some believe that quantum chemistry will be one of the first industrially viable applications of quantum computing [118]. This speculation led researchers working in industry to publish a series of papers which, for the first time, aimed to realistically estimate the resources required to solve an important classically intractable instance of the molecular electronic structure problem on a quantum computer [281, 140, 224, 15]. The first of these papers concluded that even with an error-corrected 100 MHz digital quantum computer, the problem would remain practically intractable as a 168 spin-orbital instance would still take approximately 300,000,000 years to solve [281]. Over the course of the next year, subsequent papers, which include Chapter 8 of this thesis, introduced algorithmic improvements and tighter error bounds which reduced those estimates by more than fourteen orders of magnitude to an estimated runtime of 300 seconds [140, 224, 15].

After applying the Jordan-Wigner or Bravyi-Kitaev Hamiltonian to Eq. 1.5, one obtains a spin Hamiltonian containing $\mathcal{O}\left(N^4\right)$ terms having locality $\mathcal{O}\left(N\right)$ or $\mathcal{O}\left(\log N\right)$, respectively [145, 243]. The typical approach to this algorithm involves three additional steps. First, one must prepare an initial ansatz state $|\psi\rangle$ which has polynomially bounded overlap with the ground state of the exact Hamiltonian, $|0\rangle$; i.e., $|\langle\psi|0\rangle|^2 \in \Omega\left(\text{poly}\left(N^{-1}\right)\right)$ [11]. One way to accomplish this would be to use adiabatic state preparation, such as in Chapter 7.

The next step is to evolve the ansatz under the molecular propagator $U(t)$. Since $H$ and $U(t)$ share eigenstates, this operator induces phases on the eigenstates of $H$ which encode the eigenvalues of $H$. Specifically, when applied to $|\psi\rangle$,

$$U(t)|\psi\rangle = e^{-iHt}|\psi\rangle = \sum_{j=0}^{2^{N-1}} e^{i2\pi\phi_j}|j\rangle\langle j|\psi\rangle = \sum_{j=0}^{2^{N-1}} e^{-iE_j t}|j\rangle\langle j|\psi\rangle \tag{1.9}$$

where $H|j\rangle = E_j|j\rangle$. Therefore, measurement of the phase projects us to an eigenstate of the Hamiltonian and if we have sufficient overlap with the ground state, we will measure the ground state energy with bounded probability. The typical strategy to measure this phase was introduced by Kitaev in [168] and elaborated on for the chemistry problem in [285, 284]. The circuit for phase estimation is shown below in Figure 1.1.

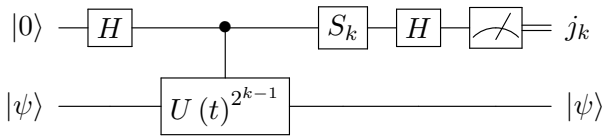



Figure 1.1: The generic circuit for iterative phase estimation of bit $k$.

The general idea of this circuit is that we want to read out the phase accumulated on the eigenstate of interest bit by bit to $L$ bits [285]. The readout values from Figure 1.1 gives the fractional binary expansion of $\phi$,

$$\phi = 0.j_1 j_2 \cdots j_L = \left(\frac{j_1}{2^1}\right) + \left(\frac{j_2}{2^2}\right) + \cdots + \left(\frac{j_L}{2^L}\right). \tag{1.10}$$

The energy of the measured eigenstate is given from the phase as $E = -2\pi\phi/t$. In general, we must choose $t$ so as to avoid aliasing the large eigenvalues, i.e. so that $t < 2\pi/(E_{\max} - E_{\min})$ [285]. In Figure 1.1, the gate $S_k$ is defined as,

$$S_k = \begin{pmatrix} 1 & 0 \\ 0 & \Phi_k \end{pmatrix} \qquad \Phi_k = \exp\left[2\pi i \sum_{l=2}^{L-k+1} \frac{j_{k+l+1}}{2^l}\right]. \tag{1.11}$$

The most costly part of this procedure is the implementation of $U(t)$ which usually relies on Trotterization. If we write our Hamiltonian as $H = \sum_\gamma H_\gamma$, then the most straightforward Trotter scheme, the first-order Trotter formula gives us

$$U(t) \approx \left( \prod_\gamma e^{-iH_\gamma t/r} \right)^r . \tag{1.12}$$

In general, the $H_\gamma$ do no commute; thus, this formula is only a good approximation when $r$, which is referred to as the Trotter number, is large. It follows that if the number of terms in the Hamiltonian is $W$, the cost of the total algorithm should scale roughly as $\mathcal{O}(Wr)$ where $r$ is the number of Trotter steps required to implement $U\left(2^{k-1}t\right)$ to one bit of accuracy. The primary contribution of Chapter 9 is to point out to the quantum information community that by using a local basis set, $W \in \mathcal{O}\left(N^2\right)$ or $W \in \mathcal{O}(N)$ when $N$ is sufficiently large. While this insight is the basis of many linear scaling methods in classical quantum chemistry, most estimates of the scaling of the quantum algorithm were based on the assumption that $W \in \mathcal{O}\left(N^4\right)$. This decreased scaling is explained in Chapter 9 as a property of the locality of physical interactions; that is, electrons always interact pairwise so the number of non-negligible terms in the Hamiltonian should only increase quadratically for a sufficiently large local basis [195].

A more nuanced issue is how one should determine the requisite Trotter number to obtain chemical precision. The initially pessimistic results of [281, 140] relied upon analytical bounds suggesting the error in Trotter formula might grow as fast $\mathcal{O}\left(N^{10}\right)$. Further analysis in [224] showed that the Trotter error in real molecules was substantially less than these bounds but still suggested that error depends substantially on the number of spin-orbitals. The primary contribution of [15], which is Chapter 8 of this thesis, is to argue that the Trotter error depends on much more than the number of spin-orbitals in a system. Instead, we use a combination of analytics and numerics to demonstrate that chemical properties

such as the largest nuclear charge in a molecule or the filling fraction of electrons in the molecular orbital diagram are more decisive for determining Trotter error. We also introduce novel methods for estimating the requisite Trotter number and provide circuits for state preparation of the molecular ansatz $|\psi\rangle$ which are asymptotically superior to results in the previously existing literature.

In Chapter 11 we describe an experimental collaboration designed to demonstrate the efficacy of the quantum phase estimation algorithm. Specifically, we use a Nitrogen-vacancy center platform to simulate Helium Hydride in a symmetry adapted configuration state function basis. Despite the limitations of existing hardware, we obtain the ground state energy of Helium Hydride in the minimal basis to $10^{-14}$ Hartree, which is far beyond the accuracy of the basis set and the requirements of chemical accuracy. This demonstration represents the most precise experimental chemistry quantum simulation ever performed.

# Part I

# Quantum Annealing

## Chapter 2

# Compiling Classical Optimization Problems for Quantum Annealing

Apart from minor modifications, this chapter originally appeared as [16]:

"Resource Efficient Gadgets for Compiling Adiabatic Quantum Optimization Problems". Ryan Babbush, Bryan O'Gorman, and Alán Aspuru-Guzik. *Annalen der Physik*. Volume 525, Number 10-11: 877-888. 2013.

## Abstract

We develop a resource efficient method by which the ground state of an arbitrary $k$-local, optimization Hamiltonian can be encoded as the ground-state of a $(k-1)$-local, optimization Hamiltonian. This result is important because adiabatic quantum algorithms are often most easily formulated using many-body interactions but experimentally available interactions are generally 2-body. In this context, the efficiency of a reduction gadget is measured by the number of ancilla qubits required as well as the amount of control precision needed to implement the resulting Hamiltonian. First, we optimize methods of applying these gadgets to obtain 2-local Hamiltonians using the least possible number of ancilla qubits. Next, we show a novel reduction gadget which minimizes control precision and a heuristic which uses this gadget to compile 3-local problems with a significant reduction in

control precision. Finally, we present numerics which indicate a substantial decrease in the resources required to implement randomly generated, 3-body optimization Hamiltonians when compared to other methods in the literature.

## 2.1 Introduction

Our group has used quantum annealing to simulate classical problems of importance in chemistry such as lattice protein folding [220, 221, 18]. During the course of this work we developed tools, explained in this paper, which are essential for practically encoding and compiling classical problems into Hamiltonians suitable for experimental implementation. The adiabatic algorithm prepares a system in the ground-state of an arbitrary Hamiltonian through adiabatic evolution from the ground-state of a trivial Hamiltonian [104]. This strategy exploits the adiabatic theorem of quantum mechanics which states that a physical system remains in its instantaneous eigenstate if a given perturbation acts on it slowly enough and if there is a gap between the eigenvalue and the rest of the Hamiltonian's spectrum [48]. In 2004, this algorithm was shown to efficiently simulate any given quantum circuit and thus to be polynomially equivalent to standard quantum computation [3]. That same year a proof by Kempe et al. demonstrated that adiabatic computation with a 2-local Hamiltonian accomplishes the same result [165].

Though unlikely to be universal for adiabatic quantum computation, the 2-local quantum Ising model with 1-local transverse field has been realized using a wide array of technologies and is known to be a stoquastic Hamiltonian for which finding the ground-state is an NP-Hard problem [23, 55, 158]. D-Wave System's current generation of quantum annealing machines are an important example of devices that implement this type of quantum

Ising model Hamiltonian[2][154]. In the last two years, a large number of academic groups have used these annealing machines to solve a diversity of practical problems from protein folding to machine learning [211, 86, 39, 147, 43]. Unfortunately, classical optimization problems are most easily formulated using many-body interactions. In order to reduce the locality of interactions from $k$-local to 2-local, one must employ the concept of a reduction gadget. A vigorous debate on the quantum nature of the D-Wave computer is currently underway [43, 278, 254]. In this paper, we avoid that discussion and focus on the issue of efficiently constructing experimentally realizable Hamiltonians.

In their 2004 proof, Kempe et al. first introduced the notion of perturbative gadgets, which use perturbation theory to reproduce the low-energy subspace of an arbitrary, QMA-Complete $k$-local Hamiltonian with a $(k-1)$-local Hamiltonian in a larger Hilbert space, expanded by the addition of ancilla qubits [165, 157]. While this tool has been essential to a number of important proofs, none have used perturbative gadgets to efficiently encode practical problems into an experimentally realizable Hamiltonian [216, 55, 38, 242]. A significant problem with using perturbative gadgets for practical encodings is that each order of perturbation theory causes an exponential increase in the control precision required to implement the resulting Hamiltonian. In this context, control precision refers to the number of distinct field values that a device must be able to resolve in order to implement the requisite interactions in a given Hamiltonian. D-Wave's newest device has 512 qubits and 4 bits of control precision, which amounts to 16 distinct values of coupler strength in both positive and negative biases. In practice, this means that problem size is often more limited by control precision than by qubits. Even for an ideal device which could implement

[2]The D-Wave hardware implements a restricted form of the 2D Ising model, known as the Chimera graph, which can encode smaller instances of the general 2D Ising model. Compiling to this specific graph is beyond the scope of the present work.

couplings to arbitrary precision (but with finite maximum field strengths) we should avoid encodings which demand high control precision. This is because the separation of energy eigenstates is inversely proportional to the control precision, and thus the adiabatic runtime increases with control precision.

Fortunately, there exist a class of non-perturbative gadgets which are significantly more efficient in terms of both ancilla and control precision [50]. We refer to these as *exact classical gadgets* because they apply only for Hamiltonians in which all many-body terms are simultaneously diagonalizable, e.g. the D-Wave final Hamiltonian. This class of Hamiltonians can encode NP-Hard problems (but not QMA-Hard unless NP = QMA) and is thus referred to as an optimization Hamiltonian [3]. Exact gadgets work by substituting the product of two qubits in a $k$-local term with an ancilla qubit and introducing a 2-local penalty function which raises the energy of any state in which the product of the original two qubits is not equal to the state of the ancilla bit. This penalty function still raises control precision considerably but we show that the effect can be partially ameliorated by a different penalty function which uses additional ancilla qubits.

The central contribution of the present work is to introduce novel techniques for efficiently applying exact classical gadgets. In order to efficiently reduce locality, one must collapse many-body terms in a systematic fashion that takes into account the appearance of specific pairs of qubits in multiple higher-order terms. For applications in which qubits are the limiting resource, we demonstrate how to map the optimal reduction to set cover and 0-1 integer linear programming (ILP) so that conventional heuristic (or exact) solvers can be leveraged to quickly find the best encoding. For control precision limited problems we formalize the optimal reduction problem and propose a greedy algorithm that significantly outperforms the status quo. Finally, we present numerics which demonstrate the signifi-

cant advantage of using these optimized gadgets and gadget application techniques over all previously mentioned reduction techniques in the literature.

## 2.2 Optimal reduction gadgets

In order to compile NP-hard optimization problems into an experimental Hamiltonian, one must encode the problem of interest into a graph of binary variables with physically realizable interactions. Perhaps the simplest model of interacting binary variables is Polynomial Unconstrained Binary Optimization (PUBO): given a pseudo-Boolean function $f : \mathbb{B}^N \to \mathbb{R}$, find an assignment $\mathbf{x} \in \mathbb{B}^N$ such that $f(\mathbf{x}) = \min\left[f\left(\mathbb{B}^N\right)\right]$, where $\mathbb{B} = \{0, 1\}$. Every pseudo-Boolean $f$ has a unique multi-linear polynomial representation

$$f(\mathbf{x}) = \sum_{S \subseteq \{1,\cdots,N\}} c_S \prod_{i \in S} x_i, \tag{2.1}$$

where $c_S \in \mathbb{R}$. From this expression we can construct an optimization Hamiltonian that embeds the energy landscape of a given PUBO in its eigenspectrum,

$$H(f) = \sum_{S \subseteq \{1,\cdots,N\}} c_S \prod_{i \in S} q_i, \tag{2.2}$$

acting on $N$ qubits, where $q_i = \frac{1}{2}\left(I^{\otimes N} - Z_i\right)$ and $Z_i$ is the Pauli matrix $\sigma^z$ acting on the $i$th qubit, i.e.

$$Z_i = I^{\otimes(i-1)} \otimes \sigma^z \otimes I^{\otimes(N-i)}, \tag{2.3}$$

where $I$ is the one-qubit identity operator. Note that while we write $H(f)$ for convenience, in practice $f$ will be specified by its coefficients $c_S$. Every element $|\mathbf{x}\rangle$ of the computational basis is an eigenstate of $H(f)$ with eigenvalue $f(\mathbf{x})$. Specifically, the ground state of $H(f)$ is spanned by the set of states $|\mathbf{x}\rangle$ such that $f(\mathbf{x}) = \min\left[f\left(\mathbb{B}^N\right)\right]$.

However, experimental interactions are typically limited to pairwise couplings between qubits, allowing Hamiltonians of the form[3]

$$H(f) = \sum_{1 \leq i \leq j \leq N} \alpha_{ij} q_i q_j, \tag{2.4}$$

where $\alpha_{ij} \in \mathbb{R}$. Such Hamiltonians correspond to a second-order pseudo-Boolean $f$,

$$f(\mathbf{x}) = \sum_{1 \leq i \leq j \leq N} \alpha_{ij} x_i x_j. \tag{2.5}$$

Thus, to encode a general instance of PUBO into an experimentally realizable Hamiltonian, one must reduce the problem to Quadratic Unconstrained Binary Optimization (QUBO), defined analogously to PUBO with the restriction that the pseudo-Boolean function to be minimized is quadratic. In practice, many common optimization problems have been reduced to PUBO in such a way that the pseudo-Boolean function to be minimized is cubic, i.e. of the form

$$f(\mathbf{x}) = \sum_{1 \leq i \leq j \leq N} \alpha_{ij} x_i x_j + \sum_{1 \leq i < j < k \leq N} \alpha_{ijk} x_i x_j x_k.$$

It is therefore desirable to have a general method for reducing a cubic function $f : \mathbb{B}^N \to \mathbb{R}$ to a quadratic function $f' : \mathbb{B}^{N'} \to \mathbb{R}$ in such a way that an assignment $\mathbf{x} \in \mathbb{B}^N$ that minimizes $f$ can be efficiently computed given an assignment $\mathbf{x}'$ that minimizes $f'$, where $N'$ is a polynomial function of $N$. One family of methods employs a set of $N' - N$ ancilla variables $\{y_1, \cdots, y_{N'-N}\} \in \mathbb{B}^{N'-N}$ such that if $(x_1, \cdots, x_N, y_1, \cdots y_{N'-N})$ minimizes $f'$, then $(x_1, \cdots, x_N)$ minimizes $f$. That is, a minimizing assignment $(x_1, \cdots, x_N)$ of $f$ is directly encoded in the $N$ computational qubits of a ground state $|x_1 \cdots x_N y_1 \cdots y_{N'-N}\rangle$ of $H(f')$. In the methods examined here, each ancilla variable corresponds to a pair of computational variables $(i, j)$ and so for convenience is denoted by $x_{ij}$ or $x_{ij}^{(m)}$.

[3]In our definition, the case in which the indices are equal is used to include 1-local terms: $q_i q_i = q_i$.

### 2.2.1 Minimal ancilla gadget

Integral to the exact gadget is the penalty function

$$s(x, y, z) = 3z + xy - 2xz - 2yz, \tag{2.6}$$

with the important property that $s(x, y, z) = 0$ if $xy = z$ and $s(x, y, z) \geq 1$ if $xy \neq z$, as shown in Table 1 [50]. While $s$ is not the only quadratic ternary pseudo-Boolean with this property, we will show that it is optimal for our purposes.

Table 2.1: Truth table for ancilla gadget

| x | y | z | $s(x, y, z)$ |
|---|---|---|---|
| 0 | 0 | 0 | 0 |
| 0 | 1 | 0 | 0 |
| 1 | 0 | 0 | 0 |
| 1 | 1 | 1 | 0 |
| 0 | 0 | 1 | 3 |
| 0 | 1 | 1 | 1 |
| 1 | 0 | 1 | 1 |
| 1 | 1 | 0 | 1 |

In our reductions, we replace a part $x_i x_j$ of a 3-local term $x_i x_j x_k$ with $x_{ij}$, where $x_{ij}$ is an ancilla variable, thereby reducing locality, while simultaneously adding the penalty function $s(x_i, x_j, x_{ij})$, scaled by an appropriate factor to ensure that the value of the reduced form is greater if $x_{ij} \neq x_i x_j$ than it is if $x_{ij} = x_i x_j$, for any assignment of the computational variables. In this way, we ensure that if an assignment of the computational and ancilla variables minimizes the reduced form, then that assignment of the computational variables also minimizes the original form. Consider the reduction

$$\alpha_{ijk} x_i x_j x_k \rightarrow \alpha_{ijk} x_{ij} x_k + (1 + |\alpha_{ijk}|)\, s(x_i, x_j, x_{ij}). \tag{2.7}$$

If $x_{ij} = x_i x_j$, then $s(x_i, x_j, x_{ij}) = 0$ and the reduced form simplifies to the unreduced form $\alpha_{ijk} x_i x_j x_k$. If $x_{ij} = 1 - x_i x_j$, then $s(x_i, x_j, 1 - x_i x_j) = 3 - 2x_i - 2x_j + 2x_i x_j$ and the reduced

form always has a greater value than it does if $x_{ij} = x_i x_j$. That is,

$$\begin{aligned} &\alpha_{ijk}\left(1 - x_i x_j\right) x_k + \left(1 + |\alpha_{ijk}|\right) s\left(x_i, x_j, 1 - x_i x_j\right) \\ &> \alpha_{ijk}\left(x_i x_j\right) x_k + \left(1 + |\alpha_{ijk}|\right) s\left(x_i, x_j, x_i x_j\right) \\ &= \alpha_{ijk} x_i x_j x_k \end{aligned} \tag{2.8}$$

for all $x_i$, $x_j$, and $x_k$. To decrease the number of ancilla variables needed to reduce many 3-local terms, it is advantageous to use the same ancilla variable $x_{ij}$ to reduce more than one 3-local term. Let $K_{ij}$ be the set of indices $k$ such that the term $x_i x_j x_k$ is reduced using the ancilla variable $x_{ij}$ corresponding to the pair of variables $\{x_i, x_j\}$. Each non-zero 3-local term is reduced using exactly one ancilla, and so we must choose $\{K_{ij}\}$ such that for each $\alpha_{ijk} \neq 0$, there is exactly one pair of indices $\{w, v\}$ with $\{w, v, K_{wv}\} = \{i, j, k\}$[4]. Then the entire set of 3-local terms can be reduced by

$$\begin{aligned} &\sum_{1 \leq i < j < k \leq N} \alpha_{ijk} x_i x_j x_k = \sum_{1 \leq i < j \leq N} \sum_{k \in K_{ij}} \alpha_{ijk} x_i x_j x_k \\ &\rightarrow \sum_{1 \leq i \leq j \leq N} \sum_{k \in K_{ij}} \left(\alpha_{ijk} x_{ij} x_k + \left(1 + |\alpha_{ijk}|\right) s\left(x_i, x_j, x_{ij}\right)\right), \end{aligned} \tag{2.9}$$

where the single term reduction in Eq. (2.7) is applied to every term in the rewritten original expression. The essential conditions (that, for any $i$ and $j$ for which an ancilla variable is used, the value of the reduced form is greater if $x_{ij} \neq x_i x_j$ than the value thereof if $x_{ij} = x_i x_j$ and in the latter case the reduced form is equal to the original form) are preserved by linearity. In Section 3, we explain a method for choosing which pair of variables to use to reduce each 3-local term (i.e. for choosing $K_{ij}$ with the constraints given) in a way that minimizes the total number of ancilla variables (the number of non-empty $K_{ij}$). In the Appendix we generalize this strategy to minimize the number of ancilla required in 4-local to 2-local reductions.

[4]Note that the indices on the coefficients are unordered, e.g. $\alpha_{ijk} = \alpha_{kji} = \alpha_{jki}$.

### 2.2.2 Minimal control precision

It is often the case that the limiting factor in encoding a PUBO instance into experimentally realizable form is the control precision rather than the number of qubits available [18]. Existing hardware is able to implement 2-local Hamiltonians of the form in Eq. (2.4) such that the coefficients are integral multiples of a fixed step size $\Delta_\alpha$ with a maximum magnitude of $N_\alpha \Delta_\alpha$, where $N_\alpha$ is the control precision. An arbitrary 2-local Hamiltonian can be made to have coefficients that are integral multiples of $\Delta_\alpha$ by dividing them all by their greatest common divisor and multiplying by $\Delta_\alpha$. The control precision needed for an arbitrary instance is thus the quotient of the greatest magnitude of the coefficients and their greatest common divisor. We assume without loss of generality that the coefficients of the PUBO to be reduced are integers and structure the reductions so that the reduced QUBO also has integral coefficients. The greatest common divisor of the coefficients of the reduced QUBO is thus one with high probability, and the control precision needed is the greatest magnitude of the coefficients. As a preliminary, we show that $s$ as defined is optimal in that the greatest coefficient (3) cannot be reduced any further.

Suppose $f(x_1, x_2, x_3)$ is a quadratic pseudo-Boolean function with integer coefficients (i.e. in the form of Eq. (2.5)) such that $f(x_1, x_2, x_3) = 0$ if $x_3 = x_1 x_2$ and is at least one otherwise. First note that $f(0,0,0) = 0$ and thus that $f(1,0,0) = \alpha_{11} = 0$ and $f(0,1,0) = \alpha_{22} = 0$. Because $f(1,1,1) = \alpha_{33} + \alpha_{12} + \alpha_{13} + \alpha_{23} = 0$, $\alpha_{33} + \alpha_{23} = -\alpha_{12} - \alpha_{13}$, and so $f(0,1,1) = \alpha_{33} + \alpha_{23} = -\alpha_{12} - \alpha_{13} \geq 1$, which implies $\alpha_{13} \leq -\alpha_{12} - 1$. Because $\alpha_{12} = f(1,1,0) \geq 1$, $\alpha_{13} \leq -2$. Finally, $f(1,0,1) = \alpha_{33} + \alpha_{13} \geq 1$ and so $\alpha_{33} \geq 1 - \alpha_{13} \geq 3$.

For each $i$ and $j$ in the reduction shown in Eq. (2.9), $(1 + |\alpha_{ijk}|)\, s(x_i, x_j, x_{ij})$ is added for each $k \in K_{ij}$, and so the coefficients in $s(x_i, x_j, x_{ij})$ are multiplied by $\sum_{k \in K_{ij}} (1 + |\alpha_{ijk}|)$.

We claim that this factor can be decreased using the reduction

$$\sum_{k\in K_{ij}} \alpha_{ijk} x_i x_j x_k \rightarrow \sum_{k\in K_{ij}} \alpha_{ijk} x_{ij} x_k + \delta_{ij} s\left(x_i, x_j, x_{ij}\right), \tag{2.10}$$

where

$$\delta_{ij} = 1 + \max\left\{ \sum_{k\in\{k\in K_{ij}|\alpha_{ijk}>0\}} \alpha_{ijk}, \sum_{k\in\{k\in K_{ij}|\alpha_{ijk}<0\}} -\alpha_{ijk} \right\},$$

for each $i$ and $j$. For all $x_i$, $x_j$, and $\{x_k|k \in K_{ij}\}$,

$$\begin{aligned} &\sum_{k\in K_{ij}} \alpha_{ijk}(1 - x_i x_j) x_k + \delta_{ij} s\left(x_i, x_j, 1 - x_i x_j\right) \\ &> \sum_{k\in K_{ij}} \alpha_{ijk}(x_i x_j) x_k + \delta_{ij} s\left(x_i, x_j, x_i x_j\right). \end{aligned} \tag{2.11}$$

That is, for any assignment of the computational variables, the value of the reduced form is greater if the ancilla variable $x_{ij} \neq x_i x_j$ than it is if $x_{ij} = x_i x_j$. The $\delta_{ij}$ given in Eq. (2.10) is optimal in the sense that it requires the least control precision of all possibilities which satisfy the appropriate conditions. Consider the reduced form

$$x_{ij} \sum_{k\in K_{ij}} \alpha_{ijk} x_k + \delta s(x_i, x_j, x_{ij}) \tag{2.12}$$

for some $\delta \in \mathbb{Z}$ to be determined. We must guarantee that

$$(1 - x_i x_j) \sum_{k\in K_{ij}} \alpha_{ijk} x_k + \delta s(x_i, x_j, 1 - x_i x_j) > x_i x_j \sum_{k\in K_{ij}} \alpha_{ijk} x_k + \delta s(x_i, x_j, x_i x_j) \tag{2.13}$$

for all $x_i$, $x_j$, and $\{x_k|k \in K_{ij}\}$. For $x_i = 1$ and $x_j = 0$ or $x_i = 0$ and $x_j = 1$, this inequality simplifies to

$$\delta > -\sum_{k\in K_{ij}} \alpha_{ijk} x_k, \tag{2.14}$$

for $x_i = x_j = 1$ it simplifies to

$$\delta > \sum_{k\in K_{ij}} \alpha_{ijk} x_k, \tag{2.15}$$

and for $x_i = x_j = 0$ it simplifies to

$$\delta > -\frac{1}{3} \sum_{k \in K_{ij}} \alpha_{ijk} x_k. \tag{2.16}$$

Eq. (2.16) is implied by Eq. (2.14) and so it is sufficient to ensure that Eq. (2.14) and Eq. (2.15) are satisfied. We see that the term $-\sum_{k \in K_{ij}} \alpha_{ijk} x_k$ is greatest when

$$x_k = \begin{cases} 1 & \text{if } \alpha_{ijk} < 0 \\ & \\ 0 & \text{if } \alpha_{ijk} > 0 \end{cases}, \tag{2.17}$$

and so if and only if

$$\delta > \sum_{k \in \{k \in K_{ij} | \alpha_{ijk} < 0\}} -\alpha_{ijk}, \tag{2.18}$$

then Eq. (2.14) is satisfied for all $\{x_k | k \in K_{ij}\}$. The term $\sum_{k \in K_{ij}} \alpha_{ijk} x_k$ is greatest under the exact opposite conditions as Eq. (2.17). Thus, if and only if

$$\delta > \sum_{k \in \{k \in K_{ij} | \alpha_{ijk} > 0\}} \alpha_{ijk} \tag{2.19}$$

then Eq. (2.15) is satisfied for all $\{x_k | k \in K_{ij}\}$. Together, Eq. (2.18) and Eq. (2.19) and imply that

$$\delta > \max \left\{ \sum_{k \in \{k \in K_{ij} | \alpha_{ijk} > 0\}} \alpha_{ijk}, \sum_{k \in \{k \in K_{ij} | \alpha_{ijk} < 0\}} -\alpha_{ijk} \right\}. \tag{2.20}$$

Note that the terms introduced in Eq. (2.10) only appear in the reduction for that pair $(i, j)$, and so the coefficient for a term therein is the coefficient in the total reduced form, with the exception of $x_i x_j$ which may also appear in the original unreduced form, which is to be addressed later. The greatest term introduced in Eq. (2.10) is $3\delta_{ij}$, which greatly increases the control precision needed.

Below, we introduce an alternative method that adds terms whose greatest coefficient is approximately a third of this. Because the complexity of the final form obscures the

simplicity of the method, we begin with a special case and extend it gradually to the general case. To reduce a single term whose coefficient is divisible by three, we introduce three ancillary bits and penalty functions:

$$\alpha_{ijk}x_ix_jx_k \to \frac{\alpha_{ijk}}{3}\left(x_{ij}^{(1)}+x_{ij}^{(2)}+x_{ij}^{(3)}\right)x_k+\left(1+\left|\frac{\alpha_{ijk}}{3}\right|\right)\sum_{m=1}^{3}s\left(x_i,x_j,x_{ij}^{(m)}\right) \tag{2.21}$$

When $x_{ij}^{(1)}=x_{ij}^{(2)}=x_{ij}^{(3)}=x_ix_j$, the reduced form simplifies to $\alpha_{ijk}x_ix_jx_k$. Otherwise, it is always greater than $\alpha_{ijk}x_ix_jx_k$, and so the reduction is valid. Furthermore, the greatest coefficient introduced is $3+|\alpha_{ijk}|$. In general however, the coefficient will not be divisible by 3. In that case, we define a new coefficient $\beta_{ijk}^{(m)}$ for each ancilla variable $x_{ij}^{(m)}$ that depends on $\alpha_{ijk}$ mod 3 such that each $\beta_{ijk}^{(m)}$ is an integer and $\sum_{m=1}\beta_{ijk}^{(m)}=\alpha_{ijk}$. This is elucidated by Table 2. We now use the reduction

Table 2.2: Integer coefficients so that $\sum_{m=1}\beta_{ijk}^{(m)}=\alpha_{ijk}$

| $\alpha_{ijk}$ mod 3 | $\beta_{ijk}^{(1)}$ | $\beta_{ijk}^{(2)}$ | $\beta_{ijk}^{(3)}$ |
|---|---|---|---|
| 0 | $\alpha_{ijk}/3$ | $\alpha_{ijk}/3$ | $\alpha_{ijk}/3$ |
| 1 | $(\alpha_{ijk}+2)/3$ | $(\alpha_{ijk}-1)/3$ | $(\alpha_{ijk}-1)/3$ |
| 2 | $(\alpha_{ijk}+1)/3$ | $(\alpha_{ijk}+1)/3$ | $(\alpha_{ijk}-2)/3$ |

$$\alpha_{ijk}x_ix_jx_k \to \left(\beta_{ijk}^{(1)}x_{ij}^{(1)}+\beta_{ijk}^{(2)}x_{ij}^{(2)}+\beta_{ijk}^{(3)}x_{ij}^{(3)}\right)x_k+\sum_{m=1}^{3}\left(1+\left|\beta_{ijk}^{(m)}\right|\right)s\left(x_i,x_j,x_{ij}^{(m)}\right). \tag{2.22}$$

If $x_{ij}^{(1)}=x_{ij}^{(2)}=x_{ij}^{(3)}=x_ix_j$, then $s(x_i,x_j,x_{ij}^{(m)})=0$ and this simplifies to $\alpha_{ijk}x_ix_jx_k$. We can rewrite the replacement terms as,

$$\sum_{m=1}^{3}\left(\beta_{ijk}^{(m)}x_{ij}^{(m)}x_k+(1+|\beta_{ijk}|)\,s\left(x_i,x_j,x_{ij}\right)\right) \tag{2.23}$$

In all cases and for each $m$

$$\beta_{ijk}^{(m)}x_{ijk}^{(m)}x_k+(1+|\beta_{ijk}|)\,s(x_i,x_j,x_{ij}^{(m)})\geq\beta_{ijk}^{(m)}x_ix_jx_k. \tag{2.24}$$

If not $x_{ij}^{(1)} = x_{ij}^{(2)} = x_{ij}^{(3)} = x_i x_j$, strict inequality holds for at least one $m$ and the replacement terms are greater than $\alpha_{ijk} x_i x_j x_k$. Here, the greatest coefficient is

$$3 + \max\left\{3\left|\beta_{ijk}^{(m)}\right|, |\alpha_{ijk}|\right\}. \tag{2.25}$$

Finally, we use the same set of ancilla variables $\left\{\alpha_{ij}^{(m)}\right\}$ to reduce all of the 3-local terms:

$$\sum_{1\le i<j<k\le N} \alpha_{ijk} x_i x_j x_k = \sum_{1\le i<j\le N} \sum_{k\in K_{ij}} \sum_{m=1}^{3} \beta_{ijk}^{(m)} x_i x_j x_k \tag{2.26}$$

$$\to \sum_{1\le i\le j\le N} \sum_{m=1}^{3} \left( \sum_{k\in K_{ij}} \beta_{ijk}^{(m)} x_{ij}^{(m)} x_k + \delta_{ij}^{(m)} s(x_i, x_j, x_{ij}^{(m)}) \right),$$

where

$$\delta_{ij}^{(m)} = 1 + \max\left\{ \sum_{k\in\{k\in K_{ij}|\beta_{ijk}^{(m)}>0\}} \beta_{ijk}^{(m)}, \sum_{k\in\{k\in K_{ij}|\beta_{ijk}^{(m)}<0\}} -\beta_{ijk}^{(m)} \right\}$$

and $K_{ij}$ is defined as above with the same constraints. In the reduced form, for every $i$, $j$, and $m$ the coefficient of $x_{ij}^{(m)}$ is $3\delta_{ij}^{(m)}$ and for every $i$ and $j$ the coefficient of $x_i x_j$ is $\sum_{m=1}^{3} \delta_{ij}^{(m)}$. The latter will be added to the coefficient $\alpha_{ij}$ of the corresponding quadratic term in the original expression. Thus the control precision needed is

$$\min_{\{K_{ij}\}} \left( \max\left\{ \max_{i,j,m}\left(3\delta_{ij}^{(m)}\right), \max_{i,j}\left|\alpha_{ij} + \sum_{m=1}^{3} \delta_{ij}^{(m)}\right| \right\} \right). \tag{2.27}$$

In Section 3 we describe a greedy algorithm to find a set of $K_{ij}$ that greatly decreases the control precision needed.

## 2.3 Efficient encoding techniques

With the exception of the 3-ancilla gadget to reduce control precision, the classical gadgets we have described have already been characterized in the literature [50]. However, knowing these formulas is not enough to efficiently encode a problem. In the following two

sections we describe how to efficiently apply these gadgets so that the resulting Hamiltonian meets the demands of available hardware. For simplicity, and because it is the most frequently encountered situation, we will focus on reductions from 3-local to 2-local. We also describe the 4-local to 2-local reduction in the Appendix of this chapter.

When working with a qubit limited encoding, the goal in applying these gadgets will be to choose the smallest set of qubit pairs that collapses all 3-local terms. We explain how to cast this problem as canonical set cover and map to 0-1 ILP so that popular heuristic or exact optimization software can be leveraged to find a set of collapsing pairs. When working with a control precision limited encoding, the goal is to choose the set of qubits for which the sum of penalty functions contains the smallest maximum coefficient. We approach this problem with a greedy algorithm but later show numerics which validate the efficiency of our technique.

### 2.3.1 Limited ancilla reduction technique

The qubit-optimized application of classical gadgets can be cast as set cover. In this context, the universe $U$ that we seek to cover is the set of 3-local terms that we must collapse. For example, $U = \{x_1x_2x_3, x_1x_4x_5, x_2x_3x_5\}$. Treating each 3-local term as a set of single qubits, we define $A$ as the union of all 2-subsets of each 3-local term. In the example given,

$$\begin{aligned} A &= \bigcup_{i=1}^{|U|} \left\{X \mid X \in 2^{U_i} \wedge |X| = 2\right\} \\ &= \{x_1x_2, x_1x_3, x_2x_3\} \cup \{x_1x_4, x_1x_5, x_4x_5\} \cup \{x_2x_3, x_2x_5, x_3x_5\} \\ &= \{x_1x_2, x_1x_3, x_1x_4, x_1x_5, x_2x_3, x_2x_5, x_3x_5, x_4x_5\}. \end{aligned} \tag{2.28}$$

Next, we construct $S$ by replacing each element $A_i$ with the union of proper supersets of $A_i$ in $U$,

$$S = \bigcup_{i=1}^{|A|} \{\{X \mid X \in U \wedge X \supsetneq A_i\}\} \tag{2.29}$$

$$= \{\{x_1x_2x_3\}, \{x_1x_2x_3\}, \{x_1x_4x_5\}, \{x_1x_4x_5\},$$

$$\{x_1x_2x_3, x_2x_3x_5\}, \{x_2x_3x_5\}, \{x_2x_3x_5\}, \{x_2x_3x_5\}\}.$$

In this way, $A$ is the set of products of pairs of qubits $x_ix_j$ that can be used in the reduction, and each element $S_i$ is the set of 3-local terms that the corresponding $A_i$ can be used to reduce. The problem is clearly set cover if we view the 3-local terms as elements (as opposed to sets themselves). Given $U$ and $S$, find the minimal covering set, i.e. $\underset{\{C|C\subseteq S\wedge\bigcup C=U\}}{\operatorname{argmin}} |C|$. In this form, the problem is easily cast as 0-1 ILP. 0-1 ILP is the problem of finding a Boolean-valued vector $v$ that minimizes the quantity $c^{\mathrm{T}}v$ subject to $Mv \geq b$. In set cover each element of $v$ is a Boolean which says whether or not to include the associated element of $S$ in the cover $C$. Thus, $c$ is a vector of ones with length equal to the cardinality of $S$ so that the cost function $c^{\mathrm{T}}v$ represents the cardinality of $C$.

The matrix $M$ multiplies $v$ to set up a system of equations which guarantees that $C$ covers $U$. Thus, the matrix element $M_{ij}$ is 1 if the $S_j$ contains the $U_i$ and 0 otherwise. Accordingly $b$ is a vector of all ones with length equal to the cardinality of $U$. Both set cover and 0-1 ILP are well known to be NP-Complete. In fact, the exact problem of cubic to quadratic polynomial binary reduction has been shown to be NP-Complete by analogy with vertex cover [50].

In Figure 2.1 we show numerics that demonstrate the efficiency of embeddings that make use of this reduction technique. For the case of 3-local to 2-local PUBO reduction, the complexity of a random problem instance is characterized by the number of logical

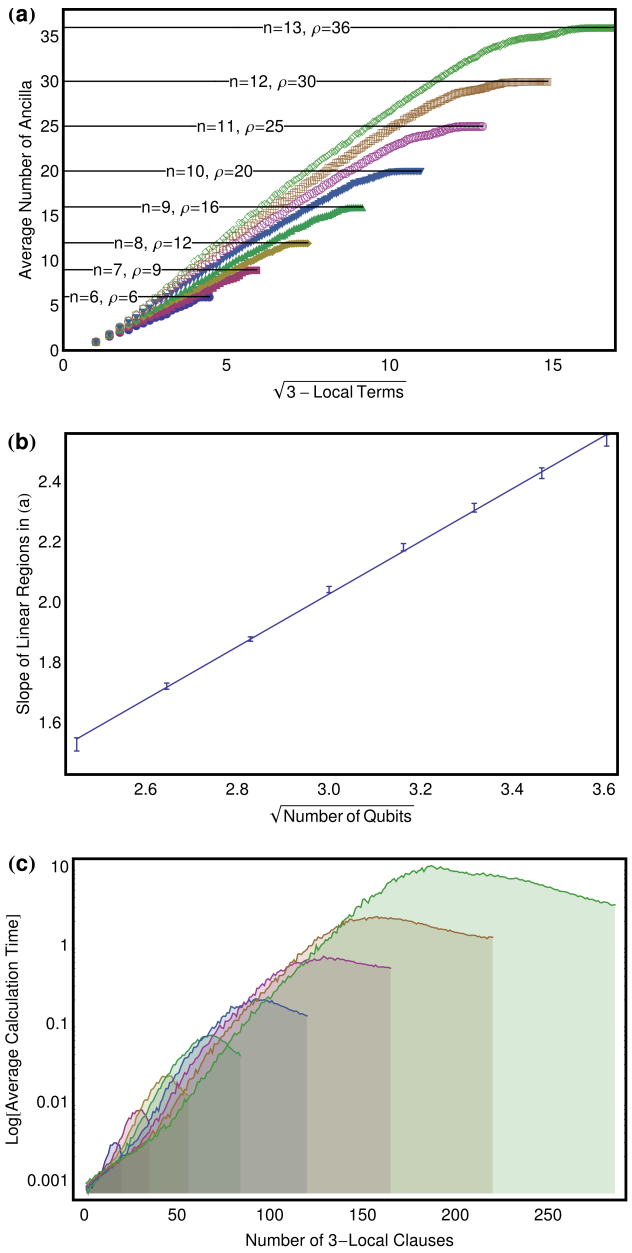



Figure 2.1: Performance of ancilla reduction scheme computed with *Mathematica 9* ILP solver. Numerics were collected on randomly generated 3-local polynomial unconstrained binary optimization (PUBO) with $n$ logical qubits and $\lambda$ 3-local clauses. For each value of $n$, data were collected from 1,000 instances at every possible value of $\lambda$, i.e. $\{\lambda \in \mathbb{Z} \mid 1 \leq \lambda \leq C_3^n\}$. Different colors indicate different values of $n$ ranging between 6 and 12 qubits. 2.1a: average number of ancilla required for reduction to 2-local versus $\sqrt{\lambda}$. 2.1b: slope of fits to linear region of aforementioned plot as a function of $\sqrt{n}$. Linear fits in top plots indicate that ancilla requirements scale as $\sqrt{n\lambda}$ until saturating at $\rho(n)$. 2.1c: semi-log plot showing average time in seconds for ILP to solve random instance.

qubits, $n$, and the number of 3-local clauses, $\lambda$. While ancilla requirements scale as $3\lambda$ for perturbative gadgets and $\lambda$ for exact gadgets without optimized application, numerics from Figure 2.1a and 2.1b indicate that our ancilla requirements scale as $\sqrt{n\lambda}$ until reaching an asymptote equal to the quarter squares function, defined as $\rho(n) = \left\lfloor \frac{(n-1)^2}{4} \right\rfloor$. A proof of this bound is shown in the Appendix. In terms of the clause to variable ratio, $r = \lambda/n$, we see that our method scales as $n\sqrt{r}$ whereas the other methods scale as $3nr$ and $nr$, respectively. Thus, we see a quadratic improvement in the number of ancilla required for a given clause to variable ratio but after a certain point, our method saturates and requires no additional ancilla, representing an undefined improvement over other methods.

Unfortunately, we should not expect to do better than a quadratic improvement for extremely large problem sizes because the constant scaling region appears to coincide with the most difficult to reduce problem instances as indicated by the computational time scaling in Figure 2.1c. In this regime, exact ILP solvers might take exponentially long to find the minimal cover. The worst case scenario is that the integrality gap of the ILP scales with the logarithm of $\lambda$, which would preclude the existence of a polynomial-time relaxation algorithm to approximate the solution beyond a logarithmic factor [9]. There does not seem to be any clear connection between the complexity of a PUBO instance and the complexity of reducing that instance to QUBO; thus, we should have no reason to suspect that on average, difficult PUBO instances will take exponential time to reduce to QUBO with ILP. However, for intractably large instances in the difficult clause to variable ratio regime, there exist greedy algorithms in the literature, for instance ReduceMin in Section 4.4 of [50] which finds the pair of indices that appears most in qubit triplets, reduces that pair in all 3-local terms, and repeats this process until all triplets are depleted.

### 2.3.2 Limited control precision reduction technique

To minimize the control precision, as expressed in Eq. (2.27), we develop a greedy algorithm which chooses the collapsing pairs, $\{K_{ij}\}$. Recall that $K_{ij}$ is the set of indices $k$ such that the term $x_i x_j x_k$ is reduced using the ancilla variable $x_{ij}$ corresponding to the pair of variables $(x_i, x_j)$. In the pseudo-code contained in Figure 2.2 we employ the convention that $K(\{i,j\}) = K_{ij}$, $\alpha(\{i,j,k\}) = \alpha_{ijk}$, and $\alpha(\{i,j\}) = \alpha_{ij}$ for ease of exposition.

The algorithm is initialized by setting $K(\{i,j\})$ to the empty set for every pair of variable indices $\{i,j\}$, and by collecting the triplet of variable indices $\{i,j,k\}$ for every 3-local term $\alpha_{ijk} x_i x_j x_k$ with a non-zero coefficient $\alpha_{ijk}$ into the set $A$. We also introduce the notation $B(a)$ for the set of three pairs of indices contained by a triplet of indices $a$, e.g. $B(\{i,j,k\}) = \{\{i,j\},\{i,k\},\{j,k\}\}$. The remainder of the algorithm consists of a procedure for choosing a 3-local term (as represented by the set of indices of its variables $d$) and a pair of variables contained therein (also represented by their indices $\Delta(d)$) with which to collapse it, which is repeated until such a choice has been made for every term that we wish to collapse. Throughout, the set $A$ contains those terms for which the decision has not been made.

The repeated procedure is as follows: first, for every 3-local term $a \in A$ for a which a pair has not been chosen with which to collapse it and for every pair therein $b \in B(a)$, the cost of collapsing the term $a$ using that pair $b$ is calculated. The cost is defined as $w(a,b) = \alpha(b) + 3 + \max\left\{\sum_{\theta\in\Theta^+} \theta, \sum_{\theta\in\Theta^-} -\theta\right\}$, where $\Theta$ is the set consisting of coefficients of terms that the pair $b$ has already been chosen to collapse and the coefficient of the current term $a$, and $\Theta^+$ and $\Theta^-$ are respectively the positive and negative elements thereof. Second, we choose a term $d$ and reduction pair $\Delta(d)$ that minimizes the cost$w$. For each term $a \in A$ we find the set of pair(s) $\Gamma(a)$ with the least cost of collapsing the term $a$.

INPUT: $N$, $\alpha : \{\{i,j,k\})|1 \leq i < j < k \leq N\} \to \mathbb{R}$
INITIALIZATION:
  **for** $1 \leq i < j \leq N$:
    $K(\{i,j\}) = \emptyset$
  $A = \{\{i,j,k\}|1 \leq i < j < k \leq N \wedge \alpha(\{i,j,k\}) \neq 0\}$
  **for** $a \in A$:
    $B(a) = \{\{p,q\}|\{p,q\} \subset a\}$
LOOP:
  **while** $|A| > 0$:
    **for** $a \in A$:
      **for** $b \in B(a)$:
        $\Theta = \{\alpha(b \cup \{k\})|k \in K(b)\} \cup \{\alpha(a)\}$
        $\Theta^+ = \{\theta|\theta \in \Theta \wedge \theta > 0\}$
        $\Theta^- = \{\theta|\theta \in \Theta \wedge \theta < 0\}$
        $$w(a,b) = \alpha(b) + 3 + \max\left\{\sum_{\theta\in\Theta^+} \theta, \sum_{\theta\in\Theta^-} -\theta\right\}$$
      $\Gamma(a) = \underset{b\in B(a)}{\operatorname{argmin}}\, w(a,b)$
      **select** $\Delta(a) \in \underset{\gamma\in\Gamma(a)}{\operatorname{argmin}}\, |\{a \in A|\gamma \subset a\}|$
    $D = \underset{a\in A}{\operatorname{argmax}}\, w(a, \Delta(a))$
    **select** $d \in D$
    $K(\Delta(d)) = K(\Delta(d)) \cup (d \setminus \Delta(d))$
    $A = A \setminus d$
OUTPUT: $K : \{\{i,j\})|1 \leq i < j \leq N\} \to 2^{\{i|1\leq i\leq N\}}$

Figure 2.2: Greedy algorithm for choosing which ancilla bits to use with each cubic term in reducing a cubic pseudo-Boolean to a quadratic one. The algorithm attempts to minimize the control precision of the reduced function. Given the function $\alpha$ that yields the coefficient of a term from the indices of its variables, the algorithm returns the function $K$ that yields the the set of indices of variables that together with the variables whose indices are passed to it form a cubic term to be reduced using the latter. See text for explanation.

Note that here we follow the convention that argmin (argmax) returns the set of arguments for which the function has its minimum (maximum) value over the specified domain, i.e. $\underset{x \in X}{\operatorname{argmin}} f(x) = \{x \in X | f(x) = \min_{x \in X} f(x)\}$.

If there is more than one such pair, we find which of those is contained in the fewest number of terms in $A$, those for which a choice has not yet been made. If there is then more than one such pair, a pair $\Delta(a)$ is chosen arbitrarily. Having found the minimum cost $w(a, \Delta(a))$ of each term $a \in A$, we find the set of terms with the minimum cost $D$ and choose one $d$ arbitrarily. Finally, we append the index in $d$ that is not in the reduction pair $\Delta(d)$ to $K(\Delta(d))$ and then remove the term $d$ from the set $A$ of terms for which a decision needs to be made. This procedure is repeated until a reduction pair has been chosen for every term, i.e. until $A$ is empty.

While we do not claim that this greedy algorithm is optimal, we present numerical evidence to show that it outperforms the default approach of selecting $K_{ij}$ in a non-systematic fashion. Figure 2.3 indicates that our technique significantly reduces the control precision cost of 3-local to 2-local reductions. For instance, with 11 qubits and 50 3-local terms, our approach requires approximately half the control precision that one would need with the default reduction strategy. Randomly choosing qubit pairs to collapse 3-local terms is the approach that many researchers (including the authors) have used in the past to encode problems into the D-Wave One device, even though the device is primarily control precision limited [221]. Our results show that the expected increase in control precision is approximately proportional to $\lambda/n$, also known as the clause to variable ratio, $r$.

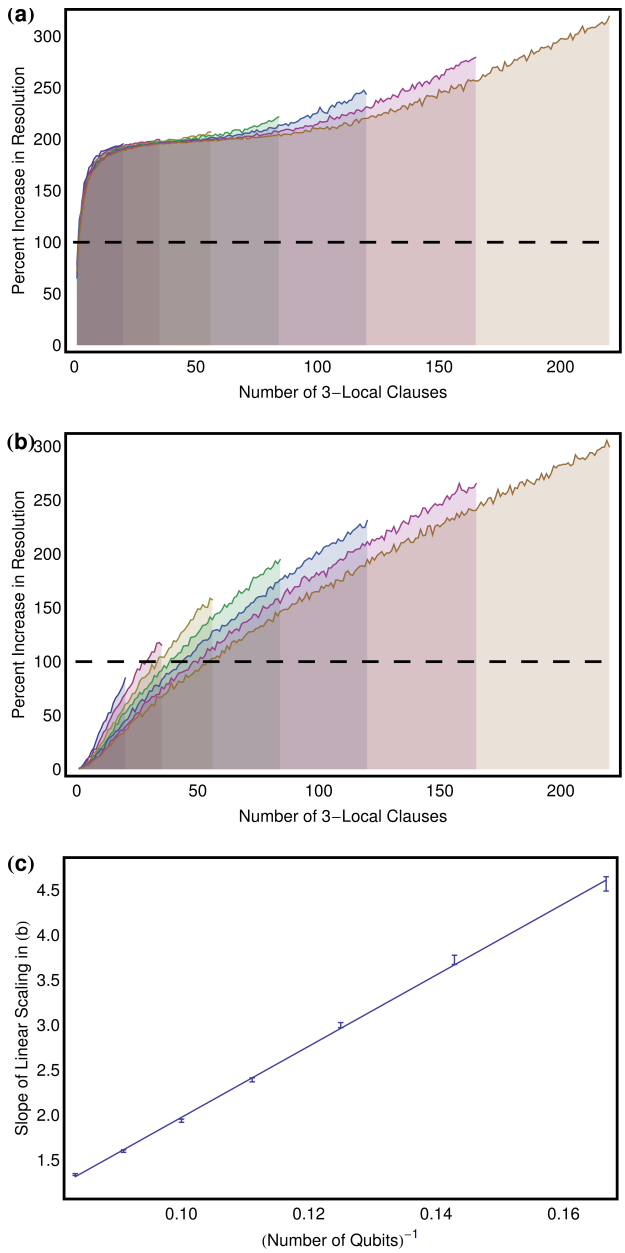



Figure 2.3: Numerics were collected on randomly generated 3-local polynomial unconstrained binary optimization (PUBO) with $n$ logical qubits, $\lambda$ 3-local clauses, and $n$ choose 2, 2-local clauses. For each value of $n$, data were collected from 1,000 instances at every possible value of $\lambda$, i.e. $\{\lambda \in \mathbb{Z} \mid 1 \leq \lambda \leq C_3^n\}$. Different colors indicate different values of $n$ ranging between 6 and 11 qubits. Integer coefficients for each term were drawn at random from a flat distribution between $-8$ and 8. With these initial coefficients, a 100% increase in control precision exhausts current D-Wave capabilities (this threshold is indicated with a dotted line). 3a: unoptimized application of reduction gadgets. 3b: application of our greedy algorithm indicating that increase in resolution is linear in $\lambda$. 3c: dependence of greedy algorithm's linear scaling in $\lambda$, suggesting control precision is proportional to $\frac{\lambda}{n}$.

## 2.4 Conclusion

In this study, we have expanded the definition of an exact classical gadget and formalized the difficult problem of efficiently applying these tools. We introduced a novel and useful form of classical gadgets that uses multiple ancilla qubits to decrease the required control precision of compiling arbitrary problems. Using this new gadget we derived Eq. (2.27), a general expression for the optimal control precision of a 3-local to 2-local reduction. While exactly solving this equation appears extremely difficult, we introduced a simple greedy algorithm which significantly outperforms the status quo. For the problem of minimizing ancilla qubit requirements during 3-local to 2-local reduction, we demonstrated how to map the problem to set cover which allowed us to find minimal ancilla encodings with the use of Integer Linear Programming. We believe that these techniques will be very useful to anyone wishing to compile classical problems into realizable Hamiltonians for adiabatic quantum computation. We are working towards applying these new techniques for protein folding and related optimization problems of interest to chemistry and biophysics.

## Acknowledgements

The authors thank David Tempel for helpful discussions. This research was sponsored by the United States Department of Defense. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the U.S. Government.

## 2.5 Appendix

### 2.5.1 Quadratization of quartic PBO by mapping to WMAXSAT

Here we show how the problem of reducing a quartic pseudo-Boolean to a quadratic one using the minimum number of ancilla bits can be recast as Weighted Max-SAT (WMAXSAT). An instance of WMAXSAT consists of a set of clauses, each of which is a disjunction of literals, and a function $w$ that assigns a non-negative weight to each clause; the problem is to find an assignment that maximizes the sum of the weights of clauses satisfied thereby.

Consider an arbitrary 4-local term $x_i x_j x_k x_l$. It can be reduced to 2-local in two ways, both of which require two ancilla bits. The first way is to use two ancilla bits that each correspond to the conjunction of two computational bits. For example, the term can be reduced using the ancilla bits $x_{ij}$ and $x_{kl}$, which entails replacing the term $x_i x_j x_k x_l$ with $x_{ij} x_{kl}$ and adding the penalty functions $s(x_i, x_j, x_{ij})$ and $s(x_k, x_l, x_{kl})$, scaled by the appropriate factor. Similarly, the term can also be reduced using $x_{ik}$ and $x_{jl}$, or $x_{il}$ and $x_{jk}$. The second way is to use an ancilla bit corresponding to the conjunction of three bits, which requires a second ancilla bit.[1] For example, the term $x_i x_j x_k x_l$ can reduced to 2-local using the ancilla bits $x_{ij}^k$ and $x_{ij}$, where $x_{ij}^k$ corresponds to the conjunction of $x_{ij}$ and $x_k$.[2] This entails replacing the term by $x_{ij}^k x_l$ and adding the penalty functions $s(x_i, x_j, x_{ij})$ and $s(x_{ij}, x_k, x_{ij}^k)$, scaled by the appropriate factor. There are twelve distinct ancilla bit pairs that can be used to reduce the term using the second way.

[1] No quadratic pseudo-Boolean $f(x, y, z, a)$ exists such that $f(x, y, z, a) = 0$ if $a = xyz$ and $f(x, y, z, a) \geq 1$ otherwise, which can be shown in a similar manner to that of the proof that the minimum coefficient in the penalty function for the conjunction of two variables is three.

[2] Accordingly, just as the indices of the ancilla bit $x_{ij}$ were unordered, i.e. $x_{ij} = x_{ji}$, so are the subscript indices of the ancilla bit $x_{ij}^k$, i.e. $x_{ij}^k = x_{ji}^k$, though the distinction between subscript and superscript indices must be made. Though in reducing a single term the choice of which pair of computational bits to use for the intermediary ancilla bit is unimportant, when reducing several the same ancilla bit may be used as an intermediary for several ancilla bits each corresponding to the conjunction of three computational bits.

Now consider a quartic pseudo-Boolean

$$f(\mathbf{x}) = \alpha_0 + \sum_{1 \le i \le N} \alpha_i x_i + \sum_{1 \le i < j \le N} \alpha_{ij} x_i x_j \tag{2.30}$$
$$+ \sum_{1 \le i < j < k \le N} \alpha_{ijk} x_i x_j x_k + \sum_{1 \le i < j < k < l \le N} \alpha_{ijkl} x_i x_j x_k x_l$$

that we would like to reduce to quadratic. Let $T_3$ and $T_4$ be sets of the sets of indices of the variables in the 3-local and 4-local terms with non-zero coefficients, respectively, i.e.

$$T_3 = \{\{i, j, k\} \subset \{1, \ldots, N\} | \alpha_{ijk} \neq 0\} \tag{2.31}$$

and

$$T_4 = \{\{i, j, k, l\} \subset \{1, \ldots, N\} | \alpha_{ijkl} \neq 0\}. \tag{2.32}$$

For each ancilla bit $x_{ij}$ that represents a conjunction of two computational bits, we introduce a Boolean variable $r_{ij} \in \{\text{TRUE}, \text{FALSE}\}$ that represents its actual use. For each triplet of computational bits $\{x_i, x_j, x_k\}$, we introduce a Boolean variable $r_{ijk} \in \{\text{TRUE}, \text{FALSE}\}$ corresponding to the use of an ancilla corresponding to their conjunction, regardless of which intermediate ancilla bit was used. While the choice of intermediate ancilla bit must be made when doing the reduction, the minimum set of ancilla bits used in a reduction cannot contain two distinct ancilla bits corresponding to the conjunction of the same three ancilla variables and so here there is no need to make the distinction. Let

$$R_2 = \left\{ r_{ij} | \{i, j\} \subset \bigcup_{t \in T_3 \cup T_4} t \right\}, \tag{2.33}$$

$$R_3 = \left\{ r_{ijk} | \{i, j, k\} \subset \bigcup_{t \in T_4} t \right\}, \tag{2.34}$$

and

$$R = R_2 \cup R_3. \tag{2.35}$$

There are three sets of clauses that must be included. First, the goal is to minimize the number of ancilla bits used in the reduction, and so for each variable representing the use of a unique ancilla bit we include the single-literal clause consisting of its negation, and assign to each such clause a weight of 1:

$$\mathcal{F}_1 = \{(\overline{r_{ij}}|\{i,j\} \in R_2\} \cup \{(\overline{r_{ijk}}|\{i,j,k\} \in R_3\} \tag{2.36}$$

and $w(C) = 1$ for every $C \in \mathcal{F}_1$. This first set consists of so-called soft clauses. The remaining two sets of clauses $\mathcal{F}_2$ and $\mathcal{F}_3$ consist of hard clauses, those that must be satisfied. This is ensured by assigning to every hard clause a weight greater than the sum of the weights of al the soft clauses. Here, we set $w(C) = |\mathcal{F}_1| + 1 = |R| + 1$ for every $C \in \mathcal{F}_2 \cup \mathcal{F}_3$. Note that $|R| \leq \binom{N}{3} + \binom{N}{2} = \frac{n(n^2-1)}{6}$.

Second, we must ensure that for each ancilla bit used that corresponds to the conjunction of three computational bits there is at least one intermediate ancilla bit that can be used in its construction, i.e.

$$(r_{ijk} \rightarrow (r_{ij} \vee r_{ik} \vee r_{jk})) \equiv (\overline{r_{ijk}} \vee r_{ij} \vee r_{ik} \vee r_{jk}). \tag{2.37}$$

Let

$$\mathcal{F}_2 = \{(\overline{r_{ijk}} \vee r_{ij} \vee r_{ik} \vee r_{jk})|\{i,j,k\} \in R_3\}. \tag{2.38}$$

Third, we must ensure that the set of ancilla bits used reduces all the cubic and quartic terms. A cubic term $x_i x_j x_k$ can be reduced using $x_{ij}$, $x_{ik}$, or $x_{jk}$, i.e. if $(r_{ij} \vee r_{ij} \vee r_{jk})$. Note that while an ancilla bit corresponding to the term itself can be used to reduce it to 1-local, that ancilla bit can only be constructed using one of the three ancilla bits mentioned, and any one of those three is sufficient to reduce the term to quadratic. A quartic term $x_i x_j x_k x_l$ can be reduced using one of twelve ancilla bits (though each requires an intermediary). These twelve can be partitioned into four triplets by the triplet of variables whose conjunction

they correspond to, i.e. by the Boolean variable that represents the use of any one. Thus the quartic term can be reduced to quadratic if $(r_{ijk} \vee r_{ijl} \vee r_{ikl} \vee r_{jkl})$. It can also be reduced using two ancilla bits that correspond to the conjunctions of disjoint pairs of computational bits, i.e. if $((r_{ij} \wedge r_{kl}) \vee (r_{ik} \wedge r_{jl}) \vee (r_{il} \wedge r_{jk}))$. These clauses must be written in conjunctive normal form:

$$\begin{aligned}&((r_{ij} \wedge r_{kl}) \vee (r_{ik} \wedge r_{jl}) \vee (r_{il} \wedge r_{jk}) \vee r_{ijk} \vee r_{ijl} \vee r_{ikl} \vee r_{jkl}) \\ &\qquad \equiv \bigwedge_{\substack{y_1 \in \{r_{ij}, r_{kl}\} \\ y_2 \in \{r_{ik}, r_{jl}\} \\ y_3 \in \{r_{il}, r_{jk}\}}} (y_1 \vee y_2 \vee y_3 \vee r_{ijk} \vee r_{ijl} \vee r_{ikl} \vee r_{jkl}).\end{aligned}$$

Let

$$\begin{aligned}\mathcal{F}_3 = &\{(r_{ij} \vee r_{ij} \vee r_{jk}) | \{i,j,k\} \in T_3\} \\ &\cup \bigcup_{\{i,j,k,l\} \in T_4} \bigcup_{\substack{y_1 \in \{r_{ij}, r_{kl}\} \\ y_2 \in \{r_{ik}, r_{jl}\} \\ y_3 \in \{r_{il}, r_{jk}\}}} (y_1 \vee y_2 \vee y_3 \vee r_{ijk} \vee r_{ijl} \vee r_{ikl} \vee r_{jkl}).\end{aligned} \tag{2.39}$$

Finally, let $\mathcal{F} = \mathcal{F}_1 + \mathcal{F}_2 + \mathcal{F}_3$. The WMAXSAT instance is specified by $\mathcal{F}$ and

$$w(C) = \begin{cases} 1 & C \in \mathcal{F}_1 \\ |R| + 1 & C \in \mathcal{F}_2 \cup F_3 \end{cases}. \tag{2.40}$$

### 2.5.2 Max number of ancilla needed to reduce a cubic pseudo-Boolean

We prove here that the minimum number of ancilla variables needed to reduce all 3-local terms over $n$ variables to 2-local is $\left\lfloor \frac{(n-1)^2}{4} \right\rfloor$, and therefore that the minimum number of ancilla variables needed to reduce any set of 3-local terms over $n$ variables is upper-bounded by the same.

The basis of the proof is Mantel's Theorem: A triangle-free graph with $n$ vertices can have at most $\left\lfloor \frac{n^2}{4} \right\rfloor$ vertices [47]. We identify a set of ancilla bits $A$ used to reduce locality

with the edge set $E(A)$ of a graph $G(A)$ whose vertices $V = \{v_i | 1 \leq i \leq N\}$ correspond to the computational variables and in which there is an edge between any two vertices $v_i$ and $v_j$ if and only if the ancilla bit $x_{ij}$ representing the conjunction of the corresponding computational bits $x_i$ and $x_j$ is used.[3] The set of ancilla bits $A$ can be used to reduce all possible 3-local terms if and only if for every set of three computational bits there is at least one ancilla bit in $A$ corresponding to the conjunction of any two. In graph-theoretic terms, $A$ can be used to reduce all 3-local terms if and only if every possible triangle in the complete graph with the same the vertex set $V$ contains at least one edge in $E(A)$, or equivalently if the complement $E^C(A)$ of $E(A)$ is triangle-free. Suppose that the set of ancilla bits $A$ reduces all 3-local terms. Then by Mantel's Theorem $|E^C(A)| \leq \lfloor \frac{N^2}{4} \rfloor$. Because $|E(A)| + |E^C(A)| = \binom{N}{2}$, this yields

$$|E(A)| = \binom{N}{2} - |E^C(A)| \geq \binom{N}{2} - \left\lfloor \frac{N^2}{4} \right\rfloor . \tag{2.41}$$

Let $N = 2m + b$, where $m = \lfloor \frac{N}{2} \rfloor \in \mathbb{Z}$ and $b = N - 2m \in \{0, 1\}$. Then

$$\begin{aligned} |E(A)| &\geq \binom{2m+b}{2} - \left\lfloor \frac{(2m+b)^2}{4} \right\rfloor \\ &= \frac{(2m+b)(2m+b-1)}{2} - \left\lfloor m^2 + mb + \frac{b^2}{4} \right\rfloor \\ &= 2m^2 + 2mb - m + \frac{b^2 - b}{2} - (m^2 + mb) \\ &= m^2 + mb - m \\ &= \left\lfloor m^2 + mb - m + \frac{b^2 - 2b + 1}{4} \right\rfloor \\ &= \left\lfloor \frac{(2m+b-1)^2}{4} \right\rfloor \\ &= \left\lfloor \frac{(N-1)^2}{4} \right\rfloor . \end{aligned} \tag{2.42}$$

[3] In reducing a cubic pseudo-Boolean to a quadratic, only ancilla bits of this type are needed.

Furthermore, by construction we show that the minimal set reaches this bound. Let $E = \{\{v_i, v_j\} | (1 \leq i < j \leq \lceil N/2 \rceil) \vee (\lceil N/2 \rceil + 1 \leq i < j \leq N\}$. That is, partition the vertices into sets of as equal size as possible and include an edge between every pair within each set. Let $N = 2m + b$ as above. The total number of edges constructed in this way is

$$\begin{aligned} \binom{\lceil N/2 \rceil}{2} + \binom{\lfloor N/2 \rfloor}{2} &= \binom{m+b}{2} + \binom{m}{2} \\ &= \frac{(m+b)(m+b-1)}{2} + \frac{m(m-1)}{2} \\ &= m^2 + mb - m \\ &= \lfloor \frac{(N-1)^2}{4} \rfloor. \end{aligned} \tag{2.43}$$

## Chapter 3

# Lattice Protein Folding Using Quantum Annealing

Apart from minor modifications, this chapter originally appeared as [18]:

"Construction of Energy Functions for Lattice Heteropolymer Models: Efficient Encodings for Constraint Satisfaction Programming and Quantum Annealing". Ryan Babbush, Alejandro Perdomo-Ortiz, Bryan O'Gorman and Alán Aspuru-Guzik. *Advances in Chemical Physics*. Volume 155, Chapter 5: 201-243. 2014.

## Abstract

Optimization problems associated with the interaction of linked particles are at the heart of polymer science, protein folding and other important problems in the physical sciences. In this chapter we explain how to recast these problems as constraint satisfaction problems such as linear programming, maximum satisfiability, and pseudo-boolean optimization. By encoding problems this way, one can leverage substantial insight and powerful solvers from the computer science community which studies constraint programming for diverse applications such as logistics, scheduling, artificial intelligence, and circuit design. We demonstrate how to constrain and embed lattice heteropolymer problems using several strategies. Each strikes a unique balance between number of constraints, complexity of constraints, and number of variables. In addition, each strategy has distinct advantages and disadvantages depending on problem size and available resources. Finally, we show

how to reduce the locality of couplings in these energy functions so they can be realized as Hamiltonians on existing quantum annealing machines.

## 3.1 Introduction

### 3.1.1 Motivation and Background

Optimization problems associated with the interaction of linked particles are ubiquitous in the physical sciences. For example, insights into a problem of biological relevance such as the protein folding problem can be obtained from trying to solve the optimization problem of finding the lowest energy configuration of a given sequence of amino acids in space [236, 90, 198, 89, 126, 245]. Among other examples of biologically relevant polymers, DNA and RNA chains also fold into complicated structures which can be challenging to predict.

The number of possible configurations (in fact, the number of local minima) for a protein with $N$ amino acids is exponential in $N$ [139]. Even the simplest model for lattice folding [177] was proved to be an NP-hard problem [27, 81]. This implies that the scaling of the worst case scenario for arbitrary protein sequences is exponential with the size of the system. This scaling imposes limitations on the exhaustive search in lattice models for proteins with as few as 36 amino acids in even the most coarse grained protein models [241].

An alternative route to exhaustive search or the development of new heuristics is to map these problems into the form of other, more general problems which have been extensively studied for decades. For instance, the NP-Complete problem known as Max-SAT has central importance to practical technologies such as artificial intelligence, circuit design, automated theorem proving, cryptography and electronic verification [130, 131, 258]. The study of this particular problem is central to computer science. There are several journals, conferences and competitions every year dedicated entirely to solving SAT prob-

lems [192]. Another widely studied constraint satisfaction problem is linear programming which has many applications including logistics scheduling, operations research, company management, and economic planning [146]. Some applications of linear programming, i.e. multi-commodity flow problems, are considered important enough that entire fields of research exist to develop specialized algorithms for their solution [102].

Once cast as one of these canonical constraint satisfaction problems one can leverage decades of progress in these fields to solve lattice heteropolymer problems. Though it has received relatively little attention until recently, the idea that constraint programming can help solve problems of this type has at least appeared in protein folding and computer science literature since [295]. Other relevant papers include [269, 84, 172, 19, 20].

Another intriguing option is to study these problems using a computer which takes advantage of quantum mechanical effects to drastically reduce the time required to solve certain problems. For combinatorial optimization problems, perhaps the most intuitive quantum computing paradigms is quantum annealing [108, 159, 239, 7, 8, 253], which is equivalent to adiabatic quantum computation [104, 103, 159] with a diagonal final Hamiltonian. In quantum annealing, the presence of quantum fluctuations (tunneling) allows the system to efficiently traverse potential energy barriers which have a tendency to trap classical optimizations algorithms.

Motivated by the experimental realization of studying biologically interesting optimization problems with quantum computation, in this contribution we present a general construction of the free-energy function for the two-dimensional lattice heteropolymer model widely used to study the dynamics of proteins. While the authors have already demonstrated some of these techniques in [220], the encoding strategies discussed here are more general and also more efficient than what we have explained previously. The reduction in

resources achieved with these methods allowed for the first experimental implementation of lattice folding on a quantum device [221] where we employed up 81 superconducting qubits to solve a six amino-acid problem in the Miyazawa-Jernigan (MJ) model [199].

The goal of this review is to explain the mapping used in [221], to discuss the strengths and weaknesses of this mapping with respect to other strategies, and to demonstrate how to map the lattice heteropolymer problem into forms which can be solved by using different types of technology and algorithms. While the focus of this paper will be on lattice protein folding, the methods introduced here have very general relevance to discrete and combinatorial optimization problems in science. Whether one decides to use a classical or a quantum (annealing) device, the mappings and techniques presented here emphasize the importance of three key considerations: energy function locality, coupler/coefficient resolution, and efficiency of encoding.

In this context, the "locality" of an expression refers to the order of the largest many-body expansion term. For instance, QUBO problems, which are a binary version of the Ising model, are said to be "2-local" because QUBO expressions never contain terms with more than two variables. This is a relevant consideration because an expression which is 3-local cannot be programmed into a quantum device with only pairwise couplings. A similar consideration applies to classical solvers. Coefficient resolution refers to the ability of a quantum device or classical solver to program coupler values to the degree of precision required for the problem. Finally, the efficiency of the encoding refers to the number of bits required to encode the problem.

### 3.1.2 Overview of Mapping Procedure

The embedding strategies presented here apply to many discrete optimization problems. Mapping these problems to a constraint programming problem is a three step process. In this section we provide a brief description of the process and expand upon each step as it applies to lattice folding in later sections.

1. **Encode solution space in computational basis**

   Define a one-to-one mapping between possible valid assignments of the problem and a bit string encoding this information. Let us denote the bit string by $\boldsymbol{q} \equiv q_1 q_2 \cdots q_n$. The way information is encoded at this point can drastically alter the nature of the following three steps so one must take care to choose a mapping which will ultimately make the best use of resources; in many cases, the most compact mapping will have a high order energy function or require many ancillary bits. Regardless of how information is encoded, the bit string must uniquely enumerate each element of the low energy solution space.

2. **Constrain energy landscape with pseudo-boolean expression**

   Construct a pseudo-boolean energy function $E(\boldsymbol{q}) = E(q_1, q_2, \cdots, q_n)$ which takes $\boldsymbol{q}$ as input and correctly reproduces the relative energies in the low energy subspace of the original problem so that the optimal solution to $E(\boldsymbol{q})$ encodes the solution to the original problem. The construction of this function is not trivial and will depend largely on how information is encoded in $\boldsymbol{q}$. At this point it may be necessary to increase the dimensionality of the solution space by adding ancillary bits. In a previous contribution, we provided a specific technique to construct the energy function for particles interacting in a lattice [220]. The purpose of this contribution is to introduce

the reader to several different types of mappings which have distinct advantages or disadvantages depending on problem size, complexity and available resources.

3. **Map boolean representation to desired constraint programming**

   In most cases one can take advantage of significantly more powerful solvers by making a final transformation from pseudo-boolean function to weighted maximum satisfiability (W-SAT), integer-linear programming (ILP), or quadratic unconstrained binary optimization (QUBO). When cast as a W-SAT problem one can take advantage of both heuristic and exact W-SAT solvers which have been developed by the computer science community and tested every year in annual "SAT Competitions". When represented as an ILP problem, one can use commercial logistics scheduling software such as IBM's CPLEX. If one wishes to implement the energy expression on a quantum device it may be necessary to manipulate the energy expression so that it contains only local fields and two-body couplings. So the final step is often to reduce the dimensionality of the pseudo-boolean expression to 2-local so that the problem can be implemented as QUBO on currently existing architectures for adiabatic quantum computing as was done in [221].

## 3.2 The "Turn" Encoding of Self-Avoiding Walks

### 3.2.1 Embedding physical structure

Let us use the term "fold" to denote a particular self-avoiding walk (SAW) assumed by the ordered chain of beads or "amino acids" on a square lattice. These configurations include amino acid chains that might intersect at different points due to amino acids occupying the same lattice sites. Even though overlapping folds will exist in the solution space of our

problem, these folds are unphysical and therefore we need to construct energy functions to penalize such configurations. Such functions will be discussed in detail below.

A fold of an $N$ amino acid protein is represented in what we refer to as the "turn" mapping by a series of $N-1$ turns. We use this name to distinguish the encoding from other (spatial) representations which encode the possible folds by explicitly encoding the grid location of each amino acid. The square lattice spatial representation discussed in [220] has the advantage of being general for the problem of $N$ particles interacting in a lattice (which need not be connected) but we can do much better in terms of the number of variables needed; bit efficiency is the main advantage of the turn mapping.

In the turn mapping, one saves bits by taking advantage of the connectivity of a valid SAW to store information about where each amino acid is relative to the previous amino acid instead of encoding explicit amino acid locations. Therefore, instead of encoding the positions of the $j$th amino acids in the lattice, we encode the $j$th turn taken by the $j+1$ amino acid in the chain. For pedagogical purposes, we concentrate on the case of a two-dimensional ($2D$) lattice SAW; the extension to a three-dimensional lattice requires a straightforward extension of the same techniques described here for the $2D$ case.

Because the location of an amino acid in the turn mapping is specified by its location relative to the previous acid in the primary sequence, the solution space consists only of paths, or "worms", embedded in the lattice. The resulting energy function is invariant under translational, rotation and reflection with respect to the embedding in physical space as long as the local structure of the relative locations is kept intact. More specifically, each of the $N-1$ turns in $2D$ space requires two bits so that each of the four directions (up, down, left, and right) has a unique representation. This assumes a rectilinear lattice, but the method is equally valid, though with slight modification, for other lattices, e.g.

triangular. The convention or "compass" used in this paper is presented in the upper-left part of Fig. 3.1. Furthermore, we can fix the first three bits to obtain only solutions which are rotationally invariant. Under this convention, the bit-string $\boldsymbol{q}$ is written as,

$$\boldsymbol{q} = 01 \underbrace{0q_1}_{turn2} \underbrace{q_2 q_3}_{turn3} \cdots \underbrace{q_{2(N-1)-4} q_{2(N-1)-3}}_{turn(N-1)} \tag{3.1}$$

We have chosen to fix the first three bits as 010 so that the walk always turns first to the right and then either right or down. This does not affect the structure of the solution space and leaves only $N-2$ turns to be specified; an example is provided in Eq. 3.1. Since every turn requires 2 bits, the turn mapping requires only $2(N-2)-1 = 2N-5$ bits to represent a fold. This can be compared with the $(2N-4)\log_2 N$ required for the spatial mapping in [220]. To clearly demonstrate how this mapping works, an example of the turn encoding for a short SAW is shown below in Fig. 3.1.

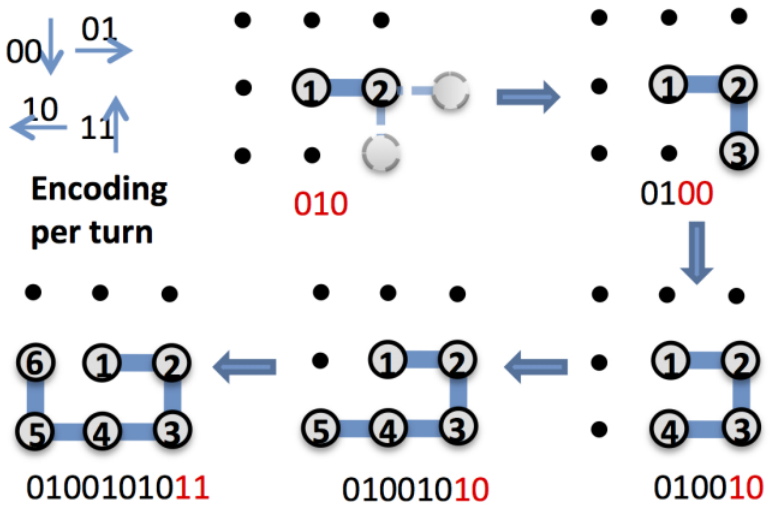



Figure 3.1: Step-by-step construction of the binary representation of a particular six unit lattice protein in the turn encoding. Two qubits per bond are needed and the turn "compass" (bond directions) are denoted as "00" (downwards), "01" (rightwards), "10" (left), and "11" (upwards). This image has been reproduced with permission from [221].

### 3.2.2 "Turn ancilla" construction of $E(\boldsymbol{q})$

Now that we have a mapping function which translates a length $2N-5$ bit-string into a specific fold in the 2D lattice we can construct $E(\boldsymbol{q})$ as a function of these binary variables. For the case of lattice folding, we need to penalize folds where two amino acids overlap, i.e. the chain must be self-avoiding. This penalty will be given by the energy function, $E_{overlap}(\boldsymbol{q})$, which returns an extremely high value if and only if amino acids overlap. While it is possible to construct a single function $E_{overlap}(\boldsymbol{q})$ which penalizes all potential overlaps, we will show that less ancillary bits are needed if we introduce the function $E_{back}(\boldsymbol{q})$ which penalizes the special case of overlaps that happen because the chain went directly backwards on itself. In this scheme, $E_{overlap}(\boldsymbol{q})$ will apply to all other potential overlaps.

Finally, we must consider the interaction energy among the different amino acids. This will ultimately determine the structure of the lowest energy fold. The energy given by the pairwise interaction of beads in our chain will be given by $E_{pair}(\boldsymbol{q})$. In some lattice protein models such as the Hydrophobic-Polar (HP) protein folding model, there is only one stabilizing interaction; however, the construction we present here applies for an arbitrary interaction matrix among the different amino acids (or particles to be even more general). One of the advantages of the turn representation over the spatial representation is that we do not need to worry about having the amino acids linked in the right order (primary sequence), since this is guaranteed by design. The construction of the energy function,

$$E(\boldsymbol{q}) = E_{back}(\boldsymbol{q}) + E_{overlap}(\boldsymbol{q}) + E_{pair}(\boldsymbol{q}), \tag{3.2}$$

involves a series of intermediate steps which we outline next.

**Construction of $E_{back}(\boldsymbol{q})$**

In order to have a valid SAW we need to guarantee that our "worm" does not turn left and then immediately turn right or vice versa or turn up and then immediately turn down or vice versa. In order to program this constraint into the energy function we will introduce several simple logic circuits. Looking at the compass provided in Fig. 3.1 it should be clear the circuits in Figs. 3.2-3.5 return TRUE if and only if a particular turn (encoded $q_1 q_2$) went right, left, up, or down respectively.

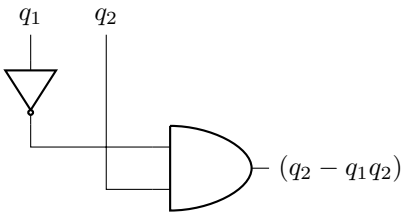



Figure 3.2: A logical circuit representing "right" consisting of a NOT gate after the first bit and an AND gate. Evaluates to TRUE if and only if $q_1, q_2 = 0, 1$.

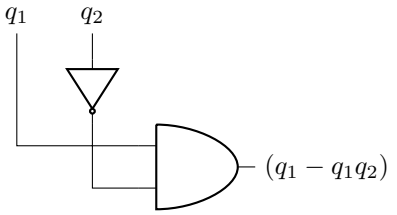



Figure 3.3: A logical circuit representing "left" consisting of a NOT gate after the second bit and an AND gate. Evaluates to TRUE if and only if $q_1, q_2 = 1, 0$.

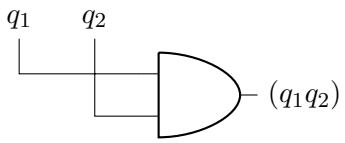



Figure 3.4: A logical circuit representing "up". Only TRUE if $q_1, q_2 = 1, 1$.

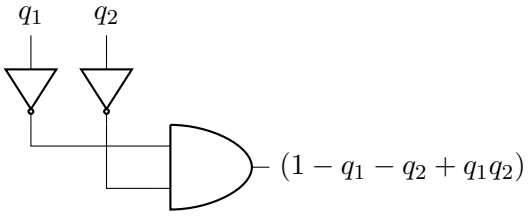



Figure 3.5: A logical circuit representing "down". Only TRUE if $q_1, q_2 = 0, 0$.

Using these circuits we can generalize the concept of "up", "down", "left" and "right" functions to precise directional strings. In two dimensions (as prescribed by Fig. 3.1), we have the functions for the $j$th turn,

$$d^j_{x+} = q_{2j}(1 - q_{2j-1}) = q_{2j} - q_{2j}q_{2j-1} \tag{3.3}$$

$$d^j_{x-} = (1 - q_{2j})q_{2j-1} = q_{2j-1} - q_{2j}q_{2j-1} \tag{3.4}$$

$$d^j_{y+} = q_{2j}q_{2i-1} \tag{3.5}$$

$$d^j_{y-} = (1 - q_{2j})(1 - q_{2j-1}) = 1 - q_{2j} - q_{2j-1} + q_{2j}q_{2j-1}, \tag{3.6}$$

which evaluate to TRUE if and only if the $j$th turn is to be right, left, up or down respectively. Having defined these circuits we can construct a more complicated circuit which takes two turns (the 4 bits $q_i q_{i+1} q_{i+2} q_{i+3}$) as input and returns TRUE if and only if the second turn went backwards, i.e. $\left(d^j_{x+} \wedge d^{j+1}_{x-}\right) \vee \left(d^j_{x-} \wedge d^{j+1}_{x+}\right) \vee \left(d^j_{y+} \wedge d^{j+1}_{y-}\right) \vee \left(d^j_{y-} \wedge d^{j+1}_{y+}\right)$. An example of these conjunctions, $\left(d^j_{x+} \wedge d^{j+1}_{x-}\right)$ is shown in Fig. 3.6.

The other three conjunctions are also trivially constructed by combining the appropriate circuits using AND gates which simply multiply together the directional strings. The utility of these circuits is that they produce terms in a pseudo-boolean function. Specifically

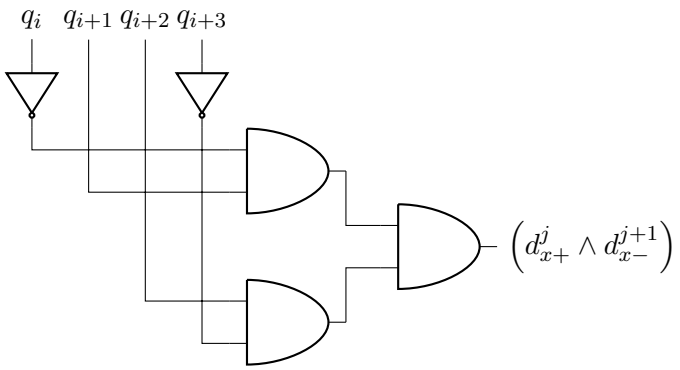



Figure 3.6: A logical circuit which returns TRUE if and only if $\left(d^j_{x+} \wedge d^{j+1}_{x-}\right)$, i.e. the turn sequence $q_iq_{i+1}q_{i+2}q_{i+3} = 0110$, meaning it went right and then left.

we get the terms,

$$\left(d^j_{x+} \wedge d^{j+1}_{x-}\right) = q_{i+1}q_{i+2} - q_iq_{i+1}q_{i+2} - q_{i+1}q_{i+2}q_{i+3} + q_iq_{i+1}q_{i+2}q_{i+3} \tag{3.7}$$

$$\left(d^j_{x-} \wedge d^{j+1}_{x+}\right) = q_iq_{i+3} - q_iq_{i+1}q_{i+3} - q_iq_{i+2}q_{i+3} + q_iq_{i+1}q_{i+2}q_{i+3} \tag{3.8}$$

$$\left(d^j_{y+} \wedge d^{j+1}_{y-}\right) = q_iq_{i+1} - q_iq_{i+1}q_{i+2} - q_iq_{i+1}q_{i+3} + q_iq_{i+1}q_{i+2}q_{i+3} \tag{3.9}$$

$$\left(d^j_{y-} \wedge d^{j+1}_{y+}\right) = q_{i+2}q_{i+3} - q_iq_{i+2}q_{i+3} - q_{i+1}q_{i+2}q_{i+3} + q_iq_{i+1}q_{i+2}q_{i+3}. \tag{3.10}$$

It might seem logical to finish this circuit by combining all four backwards overlap circuits with OR gates; however, this is not an advisable strategy as it is sure to produce many high ordered terms. Because exactly one or none of these circuits will be TRUE we can accomplish the same result by summing the four circuits. Accordingly, for the two turns $q_iq_{i+1}q_{i+2}q_{i+3}$ the pseudo-boolean expression,

$$\left(d^j_{x+} \wedge d^{j+1}_{x-}\right) + \left(d^j_{x-} \wedge d^{j+1}_{x+}\right) + \left(d^j_{y+} \wedge d^{j+1}_{y-}\right) + \left(d^j_{y-} \wedge d^{j+1}_{y+}\right) \tag{3.11}$$

evaluates to TRUE if and only if $q_iq_{i+1}q_{i+2}q_{i+3}$ represents a backwards turn and evaluates to FALSE otherwise. Our goal is to construct a pseudo-boolean expression which returns a penalty whenever a backwards turn is made; therefore we must multiply this expression by

a constant to be determined later known as $\lambda_{overlap}$. After substituting Eqs. 3.7-3.10 into Eq. 3.11, factoring the terms, and adding in $\lambda_{overlap}$ we can write,

$$E_{back}\left(q_i q_{i+1} q_{i+2} q_{i+3}\right) = \lambda_{overlap}\left(2q_i q_{i+2} - q_i - q_{i+2}\right)\left(2q_{i+1} q_{i+3} - q_{i+1} - q_{i+3}\right). \quad (3.12)$$

To construct the entire $E_{back}\left(\boldsymbol{q}\right)$ we need to sum together bits from each pair of adjacent turns. Keeping in mind that we fix the first three bits at 010, we write the final expression for $E_{back}(\boldsymbol{q})$ as,

$$\begin{aligned} E_{back}(\boldsymbol{q}) &= \lambda_{overlap}\left(q_1 q_2 + q_2 q_3 - 2q_1 q_2 q_3\right) \\ &+ \lambda_{overlap}\textstyle\sum_{i=2}^{2N-8}\left(2q_i q_{i+2} - q_i - q_{i+2}\right)\left(2q_{i+1} q_{i+3} - q_{i+1} - q_{i+3}\right). \end{aligned} \quad (3.13)$$

In this expression, the first three terms come from ensuring that the second turn (which begins with a bit fixed at 0) does not overlap with the third turn. Notice that in this expression, the first physical bit with an unknown value is labeled "$q_1$" despite the fact that the first three information bits are fixed at 010. This formalism will be consistent throughout our review.

It is important to point out that while the decision to use a separate $E_{back}(\boldsymbol{q})$ instead of a more general $E_{overlap}\left(\boldsymbol{q}\right)$ has the disadvantage of introducing 3 and 4-local terms, it has the advantage of construction without any ancillary bits. Furthermore, even if one needs an entirely 2-local expression this strategy may still be preferable because the same reductions needed to collapse this expression to 2-local will be needed in collapsing the pairwise energy function to 2-local by construction. For more on reductions, see Sec. 3.6.

### Construction of $E_{overlap}\left(\boldsymbol{q}\right)$ with ancilla variables

The overlap energy function $E_{overlap}(\boldsymbol{q})$ penalizes configurations in which any two amino acids share the same lattice point. The penalty energy associated with any pair

of amino acids overlapping must be large enough to guarantee that it does not interfere with the spectrum of the valid configurations (we return to the topic of choosing penalty values later on). We begin by defining a function which specifies the $x$ and $y$ grid positions of each amino acid. Because the directional strings we defined earlier in Eqs. 3.7-3.10 keep track of the direction of every step we can define these functions as,

$$x_n = 1 + q_1 + \sum_{k=2}^{n-1} \left( d_{x+}^k - d_{x-}^k \right) \tag{3.14}$$

$$y_n = q_1 - 1 + \sum_{k=2}^{n-1} \left( d_{y+}^k - d_{y-}^k \right) \tag{3.15}$$

where the position of the $n$th amino acid in the sequence is a function of the preceeding $n-1$ turns iterated through with index $k$. Note that the terms in front of the sum are determined by the first three (fixed) bits: 010. With these definitions we can make an extremely useful function which will return the square of the grid distance between any two amino acids (denoted $i$ and $j$):

$$g_{ij} = (x_i - x_j)^2 + (y_i - y_j)^2 . \tag{3.16}$$

$g_{ij}$ has several extremely useful properties worth pointing out now. First, $g_{ij}$ is zero if and only if two amino acids overlap; otherwise, $g_{ij}$ is always positive. Additionally, $g_{ij}$ has the very surprising property of being natively 2-local when constructed using the compass that we defined in Fig. 3.1 (therefore the decision to encode directions in that fashion was not arbitrary). This is surprising because the directional strings are 2-local so we might naively expect something which involves the square of these to be 4-local; however this turns out not to be the case because $x_n$ and $y_n$ are 1-local by construction.

In order to use $g_{ij}$ to construct $E_{overlap}(\boldsymbol{q})$ we need a function which takes $g_{ij}$ as input and returns a penalty if and only if $g_{ij} = 0$. First, we note the bounds on $g_{ij}$,

$$0 \leq g_{ij} \leq (i-j)^2. \tag{3.17}$$

To help enforce the constraint that $g_{ij} \geq 1$, we introduce a free parameter, $\alpha_{ij}$. In the optimization literature, such a variable is called a "slack variable" and is used to convert an inequality into an equality. In our case,

$$0 \leq \alpha_{ij} \leq (i-j)^2 - 1 \tag{3.18}$$

This implies that,

$$\forall\, g_{ij} \geq 1 \,\exists\, \alpha_{ij} : (i-j)^2 - g_{ij} - \alpha_{ij} = 0. \tag{3.19}$$

Furthermore, if and only if $g_{ij} = 0$,

$$(i-j)^2 - g_{ij} - \alpha_{ij} \geq 1 \,\forall\, \alpha_{ij}. \tag{3.20}$$

In order to introduce a slack variable such as $\alpha_{ij}$ into the construction of our pseudo-boolean function we must encode it using ancilla bits. Ancilla bits are real, unconstrained bits used in the calculation which have no physical significance to the particular problem mapping (i.e. ancilla bits do not tell us anything about a particular protein fold). In using ancilla we increase the dimensionality of the solution space of our problem by introducing extra variables but gain the ability to use those bits in our energy function.

Every pair of amino acids which could possibly overlap will need unique bits to form an $\alpha$ for use in the $E_{overlap}(\boldsymbol{q})$ term corresponding to that pair. Only amino acids which are an even number of turns apart can possibly overlap and we are already preventing amino acids which are two turns apart from overlapping with $E_{back}(\boldsymbol{q})$; thus, the number of amino acid pairs which require a slack variable is calculated as,

$$\sum_{i=1}^{N-4} \sum_{j=i+4}^{N} \left[(1+i-j) \bmod 2\right]. \tag{3.21}$$

Each $\alpha_{ij}$ can be represented in binary using the corresponding ancilla bits. Using Eq. 3.18 we see that the $\alpha_{ij}$ corresponding to amino acid pair $i, j$ can be represented in $\mu_{ij}$ ancilla

bits where,

$$\mu_{ij} = \lceil 2 \log_2 (i-j) \rceil \left[ (1+i-j) \bmod 2 \right]. \tag{3.22}$$

Therefore, the total number of ancilla bits required to form $E_{overlap}(\boldsymbol{q})$ is,

$$\sum_{i=1}^{N-4} \sum_{j=i+4}^{N} \mu_{ij}. \tag{3.23}$$

Finally, we can write the formula for a given $\alpha_{ij}$ as,

$$\alpha_{ij} = \sum_{k=0}^{\mu_{ij}} q_{c_{ij}+k} 2^k \tag{3.24}$$

where $c_{ij}$ denotes a pointer to the first ancilla bit corresponding to a particular amino acid pair. For instance, if the $E_{overlap}(\boldsymbol{q})$ ancilla are in sequential order from lowest index pair to highest index pair and come immediately after the information, bits then we could write,

$$c_{ij} = \sum_{m=1}^{i} \left( \sum_{n=m+4}^{N} \mu_{mn} \right) - \sum_{n=j}^{N} \mu_{in}. \tag{3.25}$$

However, there are still several problems we must address before we can construct $E_{overlap}(\boldsymbol{q})$. To begin with, we originally wanted an $\alpha_{ij}$ which was specifically restricted to the domain given in Eq. 3.18 but since we cannot constrain the physical bits in any fashion, Eq. 3.22 and Eq. 3.24 suggest that our slack variable is actually in the domain given by,

$$0 \leq \alpha_{ij} \leq 2^{\mu_{ij}} - 1. \tag{3.26}$$

We should adjust Eq. 3.19 and Eq. 3.20 so that,

$$\forall\, g_{ij} \geq 1 \,\exists\, \alpha_{ij} : 2^{\mu_{ij}} - g_{ij} - \alpha_{ij} = 0. \tag{3.27}$$

Furthermore, if and only if $g_{ij} = 0$,

$$2^{\mu_{ij}} - g_{ij} - \alpha_{ij} \geq 1 \,\forall\, \alpha_{ij}. \tag{3.28}$$

Finally, there is the question of how to guarantee that $\alpha_{ij}$ is the particular $\alpha_{ij}$ that gives 0 in Eq. 3.27 whenever $g_{ij} \geq 1$. Even though there exist $\alpha_{ij}$ such that Eq. 3.27 evaluates to 0, it is also possible to have $\alpha_{ij}$ such that Eq. 3.27 evaluates to a negative value. Negative values would incentivize overlaps instead of penalizing them so to ensure that the lowest energy solution always has $E_{overlap}(\boldsymbol{q}) = 0$ we square the expression to obtain the following formula,

$$\gamma_{ij} = \lambda_{overlap} \left[2^{\mu_{ij}} - g_{ij} - \alpha_{ij}\right]^2. \tag{3.29}$$

The expression $\gamma_{ij}$ is effective for our purposes because $\alpha_{ij}$'s restricted domain given by Eq. 3.26, promises that $\gamma_{ij}$ can only equal zero if $g_{ij} \geq 1$. $\gamma_{ij}$ is zero only if $g_{ij} \geq 1 \wedge \alpha_{ij} = 2^{\mu_{ij}} - g_{ij}$; thus, the goal is to make $\lambda_{overlap}$ a sufficiently large penalty that all low energy solutions must have no overlaps, i.e. $g_{ij} \geq 1$ for all $ij$, and $\alpha_{ij} = 2^{\mu_{ij}} - g_{ij}$. Finally we can write the final expression,

$$E_{overlap}(\boldsymbol{q}) = \sum_{i=1}^{N-4} \sum_{j=i+4}^{N} \left[(1 + i - j) \bmod 2\right] \gamma_{ij}. \tag{3.30}$$

Again, we include the term $[(1 + i - j) \bmod 2]$ because only amino acids that are an even number apart have the possibility of overlapping. Furthermore, because overlaps between adjacent amino acids are impossible and overlaps between amino acids two apart are prevented by $E_{back}(\boldsymbol{q})$, we start the second sum at $j = i + 4$ Accordingly, one should only create ancillary bits for pairs in which $(i - j) \bmod 2 = 0 \wedge i - j \geq 4$. It should now be clear that the reason we introduced $E_{back}(\boldsymbol{q})$ was so that we used fewer ancillary bits in this step.

### Construction of $E_{pair}(\boldsymbol{q})$ with ancilla variables

Finally, we need to construct the pairwise interaction energy function. To do this we need to make an interaction matrix, $J$, which contains all of the pairwise interactions which

lower the energy when two amino acids are adjacent on the lattice (thus all elements of $J$ are negative or zero). Note that this interaction matrix must contain many zero-valued elements as many amino acid pairs cannot possibly be adjacent. For instance, only amino acids which are at least three turns apart and an odd number of turns apart can ever be adjacent. Furthermore, depending on the interaction model many of these amino acids might not "interact"; for instance, in the HP-model only H-H pairs can interact where as in the Miyazawa-Jernigan model all amino acids can interact.

For each potential interaction, we must introduce one ancillary bit denoted $\omega_{ij}$ where $i$ and $j$ denote the amino acids involved in the interaction. $\omega_{ij}$ is essentially a switch which is only "on" without incurring an energy penalty if two amino acids are interacting (that is, if $g_{ij} = 1$). We can now write the pairwise interaction term:

$$\varphi_{ij} = \omega_{ij} J_{ij} \left(2 - g_{ij}\right) \tag{3.31}$$

This simple function does everything we need to write the pair function. Because $E_{overlap}\left(\boldsymbol{q}\right)$ ensures that $g_{ij} \geq 1$, we see that $\varphi_{ij}$ is only positive if both $J_{ij}$ and $\omega_{ij}$ are non-zero and $g_{ij}$ is greater than 2. Such solutions will never be part of the low-energy landscape for our problem because the energy could be made lower by trivially flipping the $\omega_{ij}$ ancillary bit. On the other hand, $\varphi_{ij} = J_{ij}$ if and only if $g_{ij} = 1 \wedge \omega_{ij} = 1$ which means that the pair is adjacent! Thus, the final form of $E_{pair}(\boldsymbol{q})$ is,

$$E_{pair}(\boldsymbol{q}) = \sum_{i=1}^{N-1} \sum_{j=i+3}^{N} \omega_{ij} J_{ij} \left(2 - g_{ij}\right). \tag{3.32}$$

### 3.2.3 "Turn circuit" construction of $E(\boldsymbol{q})$

The turn ancilla construction has the advantage of providing an energy expression with relatively few many-body terms but it does so at the cost of introducing ancilla bits.

If one intends to use a pseudo-boolean solver or a quantum device with adjustable many-body couplings, bit efficiency is much more important than the particular structure of the energy expression. This section explains the so-called "circuit" construction which provides optimal efficiency at the cost of introducing high ordered many-body terms. The turn circuit construction (along with reductions explained in Sec. 3.6) was used to encode problems into a quantum annealing machine in [221].

### Sum strings

The circuit construction works by keeping track of the turns in between amino acids to determine if the amino acids overlap or not. To do this we keep track of the turns in every direction using the directional strings defined in Eqs. 3.7-3.10. Using these directional strings we introduce ancillary bits referred to as "sum strings". Sum strings are strings of $\lceil \log_2(j-i) \rceil$ bits for each segment of the chain between amino acids $i$ and $j$, with $1 \leq i < j \leq N$ and $i+1 < j$. As in the case of the directional strings, we require one 'sum string" per direction per pair of amino acids to be compared. Each represents, in binary, the number of total turns in a particular direction within the segment.

As in the ancilla construction, whether or not two amino acids interact or overlap depends on the sequence of turns between them. To determine this, for each segment of the directional strings we construct a string that is the sum, in binary, of the bits between two amino acids, i.e. the total number of turns in that direction. This process is most straightforwardly described using a circuit model. Consider, a single Half-Adder gate (HA) consisting of an AND and an XOR gate, as shown in Fig. 3.7. The output of a Half-Adder can be interpreted as the two-bit sum of its two input bits. Accordingly, if we wanted to add three bits we could add two of them, and then add the resultant two-bit number to the

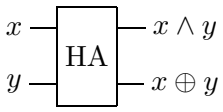



Figure 3.7: The Half-Adder gate sums two bits.

third bit, as shown in Figure 3.8.

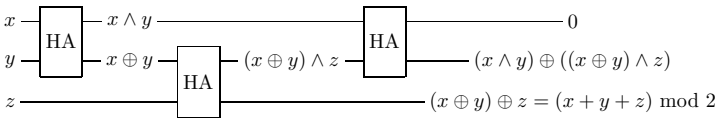



Figure 3.8: A circuit to sum three bits.

In general, to add a single bit to an $n$-bit number, we simply apply $n$ Half-Adders. First, a Half-Adder applied to the single bit and the least significant bit of the augend gives the least significant bit of the sum. Next, we use a second Half-Adder to add the carry bit of the first addition and the second least significant bit of the augend to give the second least significant bit of the sum. This process is repeated until the $(n+1)$-bit sum is computed. For an example of this see Fig. 3.9.

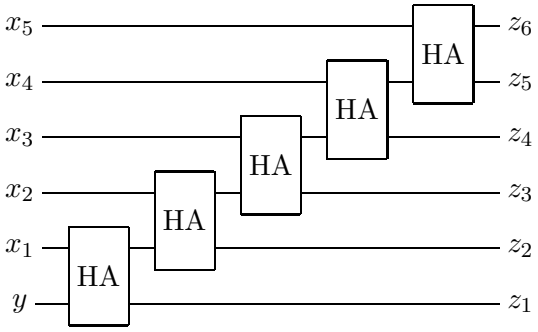



Figure 3.9: Circuit for the addition of a single bit $y$ to the 5-bit $x = x_5x_4x_3x_2x_1$ to form the 6-bit sum $x + y = z_6z_5z_4z_3z_2z_1$.

Thus, given an arbitrary number of bits we can find their sum, in binary, by successively combining the strategies shown in Fig. 3.10, i.e. first adding the first three bits (see the first

three HA gates from left to right) and then adding the next bit to the resulting three bit number which carries the previous sum. This is accomplished by the next three HA gates. From then and on, one adds a simple bit to each of the resulting $n$ bit number by using $n$ HA gates until all bits in the string are added.

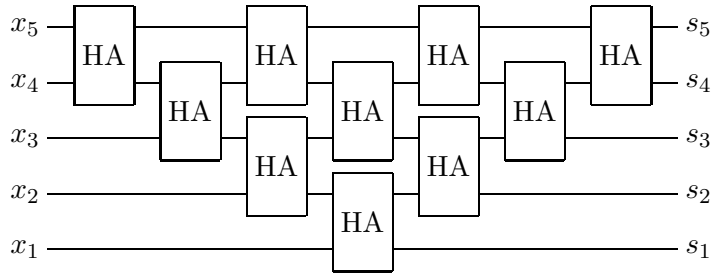



Figure 3.10: The circuit for the sum, $s_1 s_2 s_3 s_4 s_5$, of 5 bits $x_1 + x_2 + x_3 + x_4 + x_5$.

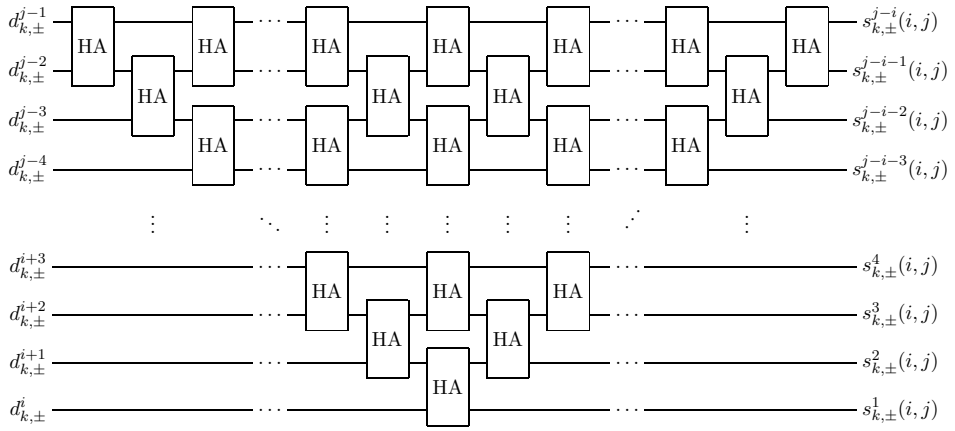



Figure 3.11: The circuit for the number $s_{k\pm}(i,j)$ of turns between amino acids $i$ and $j$ in the $\pm k$ direction.

We can use the circuit such shown in Fig. 3.11 to compute the the binary digits of a particular sum that will be very useful to us,

$$s^r_{k\pm}(i,j) = r^{\text{th}} \text{ digit of } \sum_{p=i}^{j-1} d^p_{k\pm}. \tag{3.33}$$

This sum tells us how many turns our protein has taken in the $k\pm$ direction between any two amino acids. For instance, $s^1_{x-}(3,9)$ would tell us the value of the 1st binary digit of an

integer representing the number of times that the protein turned in the negative $x$ direction (aka left) between amino acids 3 and 9. While the size of the output of the circuit given in Fig. 3.11 scales exactly with the size of the input, the maximum number of bits needed to represent the sum of a set of bits scales logarithmically; therefore, many of the bits representing higher places in the sequence are zero. Specifically, the sum of $n$ bits requires at most $\lceil \log_2 n \rceil$ bits to represent in binary.

### Construction of $E_{overlap}(\boldsymbol{q})$ with circuit

The overlap penalty should be positive if any two amino acids are at the same lattice point. For a pair $i, j$, this occurs when the number of turns between them in each direction $k\pm$ is equal to those in the opposite direction $k\mp$ or equivalently, when the bit-strings representing those numbers, $s^{j-i}_{k+} \cdots s^1_{k+}$ and $s^{j-i}_{k-} \cdots s^1_{k-}$, are the same. As discussed above, since only the first $\lceil \log_2(j-i) \rceil$ digits of $s_{k\pm}$ are non-zero, the overlap penalty function for amino acids $i$,$j$ is

$$E_{overlap}(i,j) = \prod_{k=1}^{D} \left( \prod_{r=1}^{\lceil \log(j-i) \rceil} \text{XNOR}\left(s^r_{k+}(i,j), s^r_{k-}(i,j)\right) \right), \tag{3.34}$$

where

$$\text{XNOR}(p,q) = 1 - p - q + 2pq \tag{3.35}$$

is the exclusive NOR function which evaluates to TRUE if and only if the two bits have the same value. Furthermore, we need not consider every pair of amino acids in the sequence because in order for the number of turns in opposite directions to be equal, there must be an even number of total turns. The total on-site penalty function is

$$E_{overlap} = \lambda_{overlap} \sum_{i=1}^{N-2} \left( \sum_{j=1}^{\lfloor (N-i)/2 \rfloor} E_{overlap}(i, i+2j) \right) \tag{3.36}$$

**Construction of $E_{pair}(\boldsymbol{q})$ with circuit**

To determine if a pair of amino acids is adjacent on the lattice without using ancilla bits is more involved. Two amino acids are adjacent if and only if the number of turns between them in opposite directions is the same in all but one dimension and the numbers of turns in the other dimension have a difference of one. The construction of the equality condition is the same as in as for the overlap function; to construct the latter condition, consider the set of 4 bit numbers. Note that when the first of two sequential binary is even, the Hamming distance between those bit-strings are the same except for the least significant bit, e.g. 0000 and 0001, 1000 and 1001, 1110 and 1111. On the other hand, sequential numbers for which the lesser one is odd differ in at least two places, depending on where the rightmost 0 is in the lesser number, i.e.

$$\begin{array}{rc} & 00000000000\cdots01 \\ + & **\cdots**011\cdots11 \\ \hline & **\cdots**100\cdots00 \end{array}, \tag{3.37}$$

as in 0011 and 0100, 0111 and 1000, and 1011 and 1100.

Let us use $p$ to denote the position of the rightmost 0 in the odd, lesser number of this comparison. There are three portions of the bit strings which need attention when comparing adjacency in this case. First, all digits from the least significant and up to $p$ need to be different. Second, all digits after $p$ need to be the same. Third, within each possible adjacency direction ($k+$ or $k-$) there needs to be a change from $p-1$ to $p$. Finally, all the digits from the least significant up to the $(p-2)$th digit need to be the same. Using these conditions, for both cases when the lesser number is either even or odd, results in the adjacency terms for each of the two dimensions and all of the possible amino acid pairs,

$a_k(i,j)$:

$$
\begin{aligned}
a_k(i,j) = &\left[\prod_{w\neq k}\left(\prod_{r=1}^{\lceil\log(j-i)\rceil}\text{XNOR}\left(s^r_{w+}(i,j), s^r_{w-}(i,j)\right)\right)\right] \\
&* \left[\text{XOR}\left(s^1_{k+}(i,j), s^1_{k-}(i,j)\right)\prod_{r=2}^{\lceil\log(j-i)\rceil}\text{XNOR}\left(s^r_{k+}(i,j), s^r_{k-}(i,j)\right)\right. \\
&+ \sum_{p=2}^{\lceil\log(j-i)\rceil}\left(\text{XOR}\left(s^{p-1}_{k+}(i,j), s^p_{k+}(i,j)\right)\prod_{r=1}^{p-2}\text{XNOR}\left(s^r_{k+}(i,j), s^{r+1}_{k+}(i,j)\right)\right. \\
&\left.\left.* \prod_{r=1}^{p}\text{XOR}\left(s^r_{k+}(i,j), s^r_{k-}(i,j)\right)\prod_{r=p+1}^{\lceil\log(j-i)\rceil}\text{XNOR}\left(s^r_{k+}(i,j), s^r_{k-}(i,j)\right)\right)\right].
\end{aligned}
\tag{3.38}
$$

Thus total contribution of the interaction between two amino acids to the total energy function is given by

$$E_{pair}(i,j) = J_{ij}\left[a_x(i,j) + a_y(i,j)\right], \tag{3.39}$$

where $J_{ij}$ is the adjacency matrix giving the energy of pairwise interactions that we used earlier. As was the case with the overlap penalty function, we need not consider all pairs of amino acids. In order for two amino acids to be adjacent there must be an odd number of turns between them, excluding the trivial case of amino acids that are adjacent in the primary sequence. Accordingly, the total pairwise interaction function is

$$E_{pair} = \sum_{i=1}^{N-3}\left(\sum_{j=1}^{\lceil(N-i-1)/2\rceil}E_{pair}(i, 1+i+2j)\right). \tag{3.40}$$

## 3.3 The "Diamond" Encoding of SAWs

There are many different ways in which one could encode a SAW into binary. Of all the alternatives to the "turn" encoding that we have considered, one stands out for a number of reasons: the so-called "diamond encoding" lends itself to an energy function which is natively 2-local (without any reductions) and which has a very sparse QUBO (quadratic unconstrained binary optimization) matrix. Despite the fact that the diamond encoding

requires no ancillary bits whatsoever, the encoding is still less bit-wise efficient than the "turn encoding". In the language of constraint satisfaction programming, this means that the clause:variable ratio is significantly lower when compared to the clause:variable in the turn encoding.

### 3.3.1 Embedding physical structure

The diamond encoding can be thought of as a more sophisticated version of the "spatial" encoding used in [220]. The key insight behind the diamond encoding is that if the first amino acid is fixed then each subsequent amino can only occupy a restricted set of lattice points which can be enumerated independent of any knowledge of the particular fold. To clarify this point and elucidate why we refer to this as the "diamond" encoding, see Fig. 3.12.

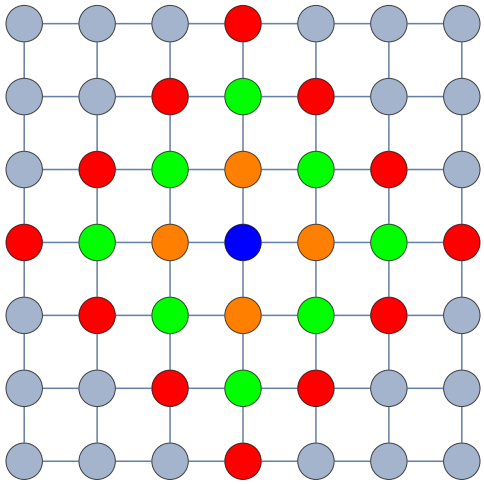


Figure 3.12: A "map" of the diamond encoding in 2D. If the first amino acid is fixed to the blue lattice point in the center then the second amino acid must be on an orange lattice point, the third must be on a green lattice point and the fourth must be on either an orange or red lattice point.

Fig. 3.12 illustrates what the "diamond" of valid lattice sites looks like for the first 4

amino acids in a SAW. In the diamond encoding each bit refers to a specific lattice site which could be occupied by an amino acid in that part of the sequence. In Fig. 3.12 we notice that the second amino acid may occupy 4 positions, the third may occupy 8 and the fourth may occupy 16. Accordingly, we need this many bits for each amino acid.

$$\boldsymbol{q} = \underbrace{q_1 q_2 q_3 q_4}_{\text{2nd acid}} \underbrace{q_5 q_6 q_7 q_8 q_9 q_{10} q_{11} q_{12}}_{\text{3rd acid}} \cdots \tag{3.41}$$

Though very straightforward to encode, this representation makes significantly less efficient use of bits than does the turn representation. However, there are a few tricks which we can use to improve the situation for this encoding. While the "diamond" of possible lattice locations for each amino acid grows quadratically with the length of the chain we can simultaneously save bits and drastically reduce the solution space without discarding the global minimum by deciding to set a hard limit on the size of the diamond. For instance, if a protein has length 100 then we would expect that the diamond for the 100th amino acid will have a radius of 99 lattice points at each corner. However, we can use the observation that proteins always fold into very compact structures to justify a substantial restriction on the solution space of our problem.

The very fact that these problems are typically called "protein folding" suggests that low energy solutions involve dense conformations. Indeed, almost all heuristic methods for folding proteins take advantage of the compact nature of low energy folds to constrain search procedures [21, 214, 246]. A large part of the reason why lattice heteropolymer problems such as protein folding are so difficult and poorly suited to heuristic algorithms is because the low energy solutions are always very compact and thus, frustrated, which makes it very unlikely that compact folds will be found efficiently via stochastic searches [88, 248, 62, 213]. Therefore, for any interesting problem its reasonable to assume that the protein will not stretch out further than a certain limit. To estimate this limit one must have familiarity

with the types of solutions expected of the particular problem. An examination of several publications holding current records for lowest energy folds in canonical problems suggests to us that for a 100 unit instance in $2D$ a reasonable cutoff radius would be around 20-30 lattice points. The cutoff radius could reasonably be made shorter for lattice models in higher dimensions as folds are expected to be even more compact on higher dimensional lattices. The number of bits required for the diamond encoding can be expected to grow cubicly up to a limit and then linearly after that limit if a cutoff is imposed. Because the number of bits required for the turn ancilla grows quadratically, for large proteins or proteins on higher dimensional lattices the diamond encoding would actually be more efficient in bits.

### 3.3.2 Natively 2-local $E(\boldsymbol{q})$

The major advantages of the diamond encoding become evident as soon as one starts to construct $E(\boldsymbol{q})$. The breakdown of the energy function looks different for the diamond encoding than for the turn encoding because the diamond encoding has different strengths and weaknesses. The first difference is that the diamond encoding will require a constraint, $E_{one}(\boldsymbol{q})$ which makes sure that each amino acid will have only one bit flipped to "on" so that each amino acid can only occupy one lattice position. Furthermore, the diamond encoding does not hard-code a primary structure constraint so we will need a term, $E_{connect}(\boldsymbol{q})$ to guarantee that each sequential amino acid is adjacent. Like the turn encoding the diamond encoding will also require $E_{overlap}(\boldsymbol{q})$ and $E_{pair}(\boldsymbol{q})$ terms. The overall energy function is,

$$E(\boldsymbol{q}) = E_{one}(\boldsymbol{q}) + E_{connect}(\boldsymbol{q}) + E_{overlap}(\boldsymbol{q}) + E_{pair}(\boldsymbol{q}). \tag{3.42}$$

**Construction of $E_{one}(\boldsymbol{q})$**

Each amino acid is encoded by flipping a bit in the part of the total bit-string sequence which represents that amino acid. Thus, we need to make sure that exactly one bit is flipped "on" for each amino acid. The most efficient way to guarantee this is the case for low energy solutions is to lower the energy whenever a bit is flipped on but introduce extremely high penalties if any two are flipped on for the same amino acid. For instance, if $\boldsymbol{q}^k$ is the binary vector which represents the $k$th amino acid and $n_k$ represents the length of this vector then we can write,

$$E_{one}(\boldsymbol{q}) = \lambda_{one} \sum_{k=2}^{N} \sum_{i=1}^{n_k-1} \sum_{j>i}^{n_k} q_i^k q_j^k. \tag{3.43}$$

$\lambda_{one}$ in Eq. 3.43 yields terms which impose a very large penalty if any two (or more) bits are flipped at once. As written, this function allows for the possibility that no bits are flipped on at once (and clearly one must be flipped on). However, the terms introduced in $E_{connect}(\boldsymbol{q})$ will guarantee that the low energy solutions all have one bit flipped on. Thus, this function only needs to make sure that no more than one bit is flipped for each amino acid.

**Construction of $E_{connect}(\boldsymbol{q})$**

To form $E_{connect}(\boldsymbol{q})$ we take a very similar approach to how we formed $E_{one}(\boldsymbol{q})$. To guarantee that the low energy solution space contains only amino acids chains which connect in the desired order we couple every bit representing amino acid $k$ to each of the $n_{k-1} \leq 4$ bits representing a lattice position adjacent to that amino acid from the previous amino acid $k-1$ and multiply by a reward as follows (using the same notation as was used in

Eq. 3.43,

$$E_{connect}\left(\boldsymbol{q}\right) = N - 2 - \lambda_{connect} \sum_{k=2}^{N} \sum_{i=1}^{n_k - 1} \sum_{j=1}^{n_{k-1}} q_i^k q_j^{k-1}. \tag{3.44}$$

Note a subtle difference between the second and third sums here is that the "-1" in the upper limit of the sum is subscripted in the latter but not in the former equation. Another important caveat is that $\lambda_{connect} << \lambda_{one}$ so that the system cannot overcome the $\lambda_{one}$ penalty by having multiple $\lambda_{connect}$ couplings. Finally we put the constant factor of $N - 2$ into the equation to adjust the energy back to zero overall for valid solutions which contain $N - 2$ connections.

**Construction of $E_{overlap}\left(\boldsymbol{q}\right)$**

It is much easier to prevent amino acids from overlapping in the diamond mapping than in the turn mapping. The only way that amino acids could overlap in the diamond mapping is for amino acids which have an even number of bonds between them to flip bits corresponding to the same lattice location. For instance, in Fig. 3.12 its clear that the fourth amino acid could overlap with second amino acid since the orange lattice points are possibilities for both amino acids. Assuming that the diamond lattice positions are encoded so that the inner diamond bits come first in the bit-string for each amino acid and that bits are enumerated in some consistent fashion (e.g. starting at the top and going clockwise around the diamond), we can write the following,

$$E_{overlap}\left(\boldsymbol{q}\right) = \lambda_{overlap} \sum_{k=2}^{N-1} \sum_{h>k}^{N} \sum_{i=1}^{n_k} \left[\left(1 + k - h\right) \bmod 2\right] q_i^k q_i^h. \tag{3.45}$$

This expression would perfectly sum over all the possible overlaps as the first two sums iterate through all possible overlapping pairs and the third sum iterates through all of the diamond points up to the last point they both share, $n_k$.

**Construction of $E_{pair}(\boldsymbol{q})$**

To form the pairwise interaction term we simply couple each bit to the possible adjacent lattice locations which could be occupied by other amino acids. The strength of the coupling will depend on the interaction matrix element between the two amino acids coupled by the term. Additionally, we note that amino acids are only able to be adjacent if there are an even number of amino acids (2 or greater) in between the two. Thus, the formula is as follows:

$$E_{pair}(\boldsymbol{q}) = \sum_{k=2}^{N-1} \sum_{h=k+2}^{N} \sum_{<ij>} J_{hk} \left[(k-h) \bmod 2\right] q_i^k q_j^h \tag{3.46}$$

where the sum over $< ij >$ is understood as a sum over bits corresponding to adjacent lattice sites. There is no straightforward way to write the function $< ij >$ in analytical terms. Nevertheless, for large problems it is trivial to write a program which iterates through bits in the second amino acid with a for-loop and evaluated the sum on those bits if the first amino acid bit and the second amino acid bit have a grid distance of 1.

## 3.4 Pseudo-boolean Function to W-SAT

In order to take advantage of state-of-the-art satisfiability (SAT) solvers to optimize our pseudo-boolean function, it is necessary to map the problem to Weighted Maximum Satisfiability (W-SAT). The most general form of the generic SAT problem is known as K-SAT. In K-SAT the problem is to find a vector of boolean valued variables which satisfies a list of clauses, each containing up to K variables, which constrain the solution. When K-SAT has a solution it is known as "satisfiable" and for K $\leq$ 2 the problem is tractable in polynomial time. However, for K $>$ 2 the problem is known to be NP-complete; in fact, 3-SAT was the first problem proved to be NP-Complete [79].

### 3.4.1 MAX-SAT and W-SAT

Maximum Satisfiability (MAX-SAT) is an alternative version of the canonical SAT problem which is relevant when K-SAT is either "unsatisfiable" or at least not known to be satisfiable. In MAX-SAT the goal is not necessarily to find the solution string which satisfies all clauses (such a solution string may not even exist); rather, the goal is to find the solution string which satisfies the maximum number of clauses.

An extension of MAX-SAT known as Weighted Maximum Satisfiability (aka W-SAT) is what will be most relevant to us. In W-SAT each clause is given a positive integer valued "weight" which is added to a sum only if the clause evaluates to FALSE. Accordingly, in W-SAT the goal is to minimize this sum rather than the total number of FALSE clauses as in canonical MAX-SAT [291, 50]. We can more succinctly state the problem as follows: given $m$ number of clauses ($y$) each with a weight of $w$, minimize

$$W = \sum_{i=1}^{m} w_i y_i \quad : \quad y_i = \begin{cases} 1 & \text{if the } i\text{th clause is FALSE} \\ 0 & \text{otherwise.} \end{cases} \tag{3.47}$$

The same approximation schemes and exact solver algorithms which work for MAX-SAT also work for W-SAT [51, 218]. In order to use these solvers one must first translate their pseudo-boolean function into a W-SAT problem articulated in what is known as Weighted Conjunctive Normal Form (WCNF). In WCNF, the W-SAT problem is stated as a list of weights followed by a clause with each clause stated as an OR statement between integers representing the index of the corresponding boolean variable in the solution vector. In WCNF, a negative integer denotes a negation. For instance the WCNF clause "4000 9 −1 82" means $x_9 \vee \neg x_1 \vee x_{82}$ with penalty of 4000 if clause evaluates to FALSE. Fig. 3.13 shows this clause as a logic circuit.

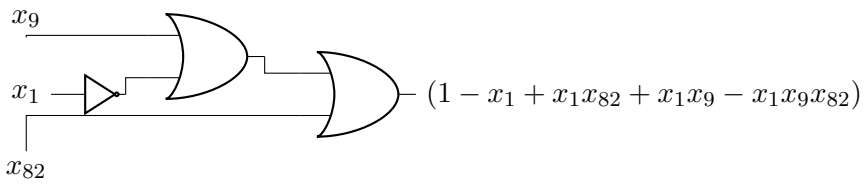



Figure 3.13: A logical circuit representation of the CNF clause: "9 −1 82"

### 3.4.2 Constructing WCNF clauses

To prepare the WCNF input file from a pseudo-boolean function one will need to write a short script which transforms each term in the pseudo-boolean function into a WCNF clause. There is more than one way to accomplish this transformation and we will only discuss one method here. For a more complete review of this topic, see [100].

It will be very useful to think of CNF clauses as logical circuits which involve only OR gates and NOT gates as in Fig. 3.13. Weights in WCNF notation always represent a positive value. Because pseudo-boolean functions are treated as cost functions to minimize and the goal of W-SAT is to minimize the sum of weights on FALSE clauses, terms in the pseudo-boolean function with a positive weight are very easy to translate in WCNF notation. To achieve this, one needs only to pass all variables in the clause through a NOT gate and then a series of OR gates (effectively making a NAND gate which takes all variables as input). This circuit is illustrated in Fig. 3.14 for the case of a 5 variables clause.

Representing a negative weighted pseudo-boolean term in CNF is less trivial but follows a simple pattern. To make the CNF clause positive (corresponding to negative boolean term) one needs to construct the same circuit as in the case when the boolean term is positive but remove one of the NOT gates. An example comprising three variables is shown in Fig. 3.15. However, this circuit alone does not accomplish our goal as it produces a 2-local term with negative weight in addition to the 3-local term with positive weight. Consequentially, after

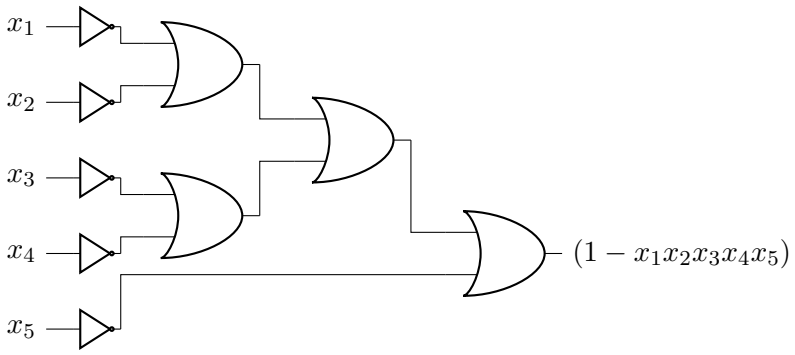



Figure 3.14: A logical circuit which shows that any pseudo-boolean term with positive weight is equivalent (up to a constant) to a CNF clause with each variable negated. The term produced here is negative because the weight is only added when the clause evaluates to FALSE.

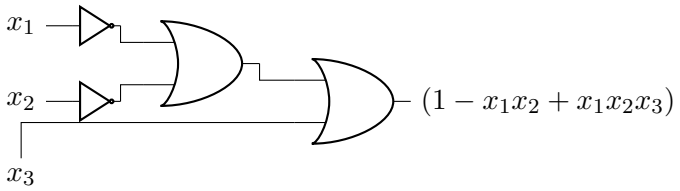



Figure 3.15: A logical circuit on three variables which gives a positive valued 3-local CNF term.

using the circuit in Fig. 3.15 to get rid of the 3-local term "$x_1x_2x_3$" we must subtract the term "$x_1x_2$" multiplied by its weight from the pseudo-boolean expression we are converting into CNF. At first glance, it is not obvious that this procedure will get us anywhere - we turned a term into CNF only to introduce a new term into the pseudo-boolean which we must convert back into CNF. However, the auxiliary terms produced by this circuit are of one degree less than number of variables in the term; thus, we can iterate this procedure until only the constant term remains. The next CNF clause (this time 2-local) is shown in Fig. 3.16.

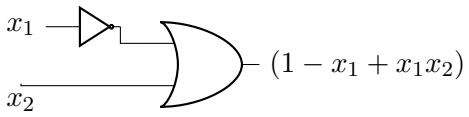



Figure 3.16: A logical circuit on three variables which gives a positive valued 2-local CNF term.

### 3.4.3 Solving SAT problems

While MAX-SAT is known to be NP-hard, there exist heuristic algorithms which are guaranteed to satisfy a fixed fraction of the clauses of the optimal solution in polynomial time. In general, oblivious local search will achieve at least an approximation ratio of $\frac{k}{k+1}$, Tabu search achieves a ratio of at least $\frac{k+1}{k+2}$ and non-oblivious local search achieves an approximation ratio of $\frac{2^k-1}{2^k}$ where $k$ is the "K" in K-SAT. For the special case of MAX-2-SAT the best possible algorithm is theoretically capable of satisfying at least $\frac{21}{22} + \epsilon \approx 0.955 + \epsilon$ [9] in polynomial time [218, 72]. Additionally, there are a great deal of exact MAX-SAT solvers which run in super-polynomial time but in many cases can find the solution to MAX-SAT in a very short amount of time, even for problems containing hundreds of variables and clauses [192, 176].

## 3.5 W-SAT to Integer-Linear Programming

Integer-Linear Programming (ILP) is a subset of linear programming problems in which some variables are restricted to integer domains. In general, ILP is an NP-Hard problem but the importance of ILP problems (particular for logistics scheduling) has produced many extremely good exponential-time exact solvers and polynomial-time heuristic solvers [291]. Pseudo-boolean optimization is an even more specific case of ILP sometimes known as 0-1 ILP where the integer variables are boolean [50]. The mapping between W-SAT and ILP

is very straightforward.

### 3.5.1 Mapping to ILP

In ILP, the goal is to minimize an objective function of integer-valued variables subject to a list of inequality constraints which must be satisfied. The inequality constraints come directly from the clauses in W-SAT. As described in Sec. 3.4.1, the logical clause from the WCNF clause "4000 9 -1 82" (which again, means $x_9 \vee \neg x_1 \vee x_{82}$ with penalty of 4000 if clause evaluates to False) can be represented as $x_9 + (1 - x_1) + x_{82} \geq 1$ s.t. $x_n \in \{0, 1\}$. In ILP, all constraints must be satisfied but in W-SAT clauses are sometimes not satisfied; to accommodate this we introduce an auxiliary binary variable, $y_1$ into the equation and get $y_1 + x_9 + (1 - x_1) + x_{82} \geq 1$. Thus, if the original equation is False, $y_1$ will have a value of True which satisfies the inequality. We can take advantage of this auxiliary variable to construct the optimization function, $W$. Since the clause in our example has a weight of 4000 we can write $W = 4000y_1$ s.t. $y_1 + x_9 + (1 - x_1) + x_{82} \geq 1$. Thus, the mapping between ILP and W-SAT is extremely trivial: all WCNF clauses are rewritten as linear equalities which are $\geq 1 - y_i$ by adding together the variables (or their negations) where $i$ is the index of the clause and the objective function is written as $W = \sum_{i=1}^{N} w_i y_i$ where $N$ is the number of clauses and $w_i$ is the weight of that clause [291].

### 3.5.2 Solving ILP problems

Commercial logistic scheduling software such as IBM ILOG CPLEX Optimization Studio (aka CPLEX) is designed to solve in integer programming, linear programming, and mixed integer-linear programming problems on a very large scale [149]. Constraint satisfaction problems which are sometimes very difficult to solve using conventional SAT techniques

can be easier to solve using ILP techniques and vice versa. In particular, SAT solvers and specialized pseudo-boolean optimizers seem to outperform ILP solvers when a problem is over-constrained [5]. On the other hand, for problems which are under-constrained and have a large number of variables ILP solvers are the natural choice. In some cases 0-1 ILP optimizers such as Pueblo will outperform both SAT solvers and commercial ILP solvers [235, 249, 190].

## 3.6 Locality Reductions

The practical ability to either exactly or approximately solve random instances of constraint satisfaction optimization such as pseudo-boolean optimization or MAX-SAT seems to depend very sensitively on the variable to clause ratio and degree of constraint expressions [218, 160, 291]. In fact, the degree of constraints determines the complexity class of certain constraint satisfaction problems; e.g. 2-SAT is proven to be in P whereas 3-SAT is in NP-Complete [79]. Clearly for instances such as this there can be no efficient method which reduces the degree of constraints. Fortunately, reducing the degree of constraints in general pseudo-boolean optimization (i.e. reducing the polynomial order of pseudo-boolean terms) can be done efficiently.

Constraint degree reduction is particularly important if we wish to solve our problem using existing architectures for adiabatic quantum computation because available devices tend to be very limited in their ability to realize arbitrary variable couplings (especially high ordered couplings). For instance, the D-Wave One device used for pseudo-boolean optimization in [221] is only able to implement 2-local qubit couplings and has limited coupler resolution. To encode functions of higher locality in such setups, we must introduce ancilla bits which replace 2-local terms to reduce locality. Because these ancilla become free

parameters of the system, it is also necessary to introduce penalty functions to account for the possibility that their value may be incorrect. All of this is accomplished with the function $E_{\wedge}\left(q_i, q_j, \tilde{q}_n; \delta_n\right)$ in Eq. 3.48 which introduces the ancillary bit $q_n$ in order to collapse the 2-local term $q_i q_j$ with energy penalty of $\delta_n$ if $q_n \neq q_i q_j$. For a further discussion, see [50, 16].

$$E_{\wedge}(q_i, q_j, \tilde{q}_n; \delta_n) = \delta_n(3\tilde{q}_n + q_i q_j - 2q_i\tilde{q}_n - 2q_j\tilde{q}_n) \tag{3.48}$$

If one desires an entirely 2-local energy function then many $E_{\wedge}\left(q_i, q_j, \tilde{q}_n; \delta_n\right)$'s may be necessary to collapse all high-local terms. For instance, consider the complete energy function for the HP model protein $HPPHP$ when coded in the turn ancilla mapping:

$$\begin{aligned}
E = & -4q_2q_6\lambda_1 + 4q_1q_3q_6\lambda_1 + 3q_6\lambda_1 + 28q_1\lambda_2 + 25q_1q_2\lambda_2 + 108q_2\lambda_2 - 56q_1q_3\lambda_2 \\
& - 50q_1q_2q_3\lambda_2 + 26q_2q_3\lambda_2 + 28q_3\lambda_2 + 24q_1q_4\lambda_2 - 16q_1q_2q_4\lambda_2 - 56q_2q_4\lambda_2 - 48q_1q_3q_4\lambda_2 \\
& + 32q_1q_2q_3q_4\lambda_2 - 18q_2q_3q_4\lambda_2 + 25q_3q_4\lambda_2 + 108q_4\lambda_2 - 56q_1q_5\lambda_2 - 48q_1q_2q_5\lambda_2 \\
& + 25q_2q_5\lambda_2 + 48q_1q_3q_5\lambda_2 - 50q_2q_3q_5\lambda_2 - 56q_3q_5\lambda_2 - 48q_1q_4q_5\lambda_2 + 32q_1q_2q_4q_5\lambda_2 \\
& - 18q_2q_4q_5\lambda_2 + 36q_2q_3q_4q_5\lambda_2 - 50q_3q_4q_5\lambda_2 + 25q_4q_5\lambda_2 + 28q_5\lambda_2 - 32q_1q_7\lambda_2 \\
& - 96q_2q_7\lambda_2 + 64q_1q_3q_7\lambda_2 - 32q_3q_7\lambda_2 + 64q_2q_4q_7\lambda_2 - 96q_4q_7\lambda_2 + 64q_1q_5q_7\lambda_2 \\
& + 64q_3q_5q_7\lambda_2 - 32q_5q_7\lambda_2 - 32q_7\lambda_2 - 16q_1q_8\lambda_2 - 48q_2q_8\lambda_2 + 32q_1q_3q_8\lambda_2 - 16q_3q_8\lambda_2 \\
& + 32q_2q_4q_8\lambda_2 - 48q_4q_8\lambda_2 + 32q_1q_5q_8\lambda_2 + 32q_3q_5q_8\lambda_2 - 16q_5q_8\lambda_2 + 64q_7q_8\lambda_2 \\
& - 32q_8\lambda_2 - 8q_1q_9\lambda_2 - 24q_2q_9\lambda_2 + 16q_1q_3q_9\lambda_2 - 8q_3q_9\lambda_2 + 16q_2q_4q_9\lambda_2 - 24q_4q_9\lambda_2 \\
& + 16q_1q_5q_9\lambda_2 + 16q_3q_5q_9\lambda_2 - 8q_5q_9\lambda_2 + 32q_7q_9\lambda_2 + 16q_8q_9\lambda_2 - 20q_9\lambda_2 - 4q_1q_{10}\lambda_2 \\
& - 12q_2q_{10}\lambda_2 + 8q_1q_3q_{10}\lambda_2 - 4q_3q_{10}\lambda_2 + 8q_2q_4q_{10}\lambda_2 - 12q_4q_{10}\lambda_2 + 8q_1q_5q_{10}\lambda_2 \\
& + 8q_3q_5q_{10}\lambda_2 - 4q_5q_{10}\lambda_2 + 16q_7q_{10}\lambda_2 + 8q_8q_{10}\lambda_2 + 4q_9q_{10}\lambda_2 - 11q_{10}\lambda_2 + 36\lambda_2.
\end{aligned} \tag{3.49}$$

In order to reduce this function to 2-local we will need to collapse some of the 2-local terms inside of the 3-local terms to a single bit. We enumerate the 3-local terms and their corresponding 2-local terms which we could use to reduce each 3-local term in Eq. 3.50.

$$\begin{pmatrix} q_1 & q_2 & q_3 \\ q_1 & q_2 & q_4 \\ q_1 & q_3 & q_4 \\ q_2 & q_3 & q_4 \\ q_1 & q_2 & q_3 \\ q_1 & q_2 & q_5 \\ q_1 & q_3 & q_5 \\ q_2 & q_3 & q_5 \\ q_1 & q_4 & q_5 \\ q_2 & q_4 & q_5 \\ q_1 & q_2 & q_4 \\ q_3 & q_4 & q_5 \\ q_2 & q_3 & q_4 \\ q_1 & q_3 & q_6 \\ q_1 & q_3 & q_7 \\ q_2 & q_4 & q_7 \\ q_1 & q_5 & q_7 \\ q_3 & q_5 & q_7 \\ q_1 & q_3 & q_8 \\ q_2 & q_4 & q_8 \\ q_1 & q_5 & q_8 \\ q_3 & q_5 & q_8 \\ q_1 & q_3 & q_9 \\ q_2 & q_4 & q_9 \\ q_1 & q_5 & q_9 \\ q_3 & q_5 & q_9 \\ q_1 & q_3 & q_{10} \\ q_2 & q_4 & q_{10} \\ q_1 & q_5 & q_{10} \\ q_3 & q_5 & q_{10} \end{pmatrix} \iff \begin{pmatrix} q_1q_2 & q_1q_3 & q_2q_3 \\ q_1q_2 & q_1q_4 & q_2q_4 \\ q_1q_3 & q_1q_4 & q_3q_4 \\ q_2q_3 & q_2q_4 & q_3q_4 \\ q_1q_2 & q_1q_3 & q_2q_3 \\ q_1q_2 & q_1q_5 & q_2q_5 \\ q_1q_3 & q_1q_5 & q_3q_5 \\ q_2q_3 & q_2q_5 & q_3q_5 \\ q_1q_4 & q_1q_5 & q_4q_5 \\ q_2q_4 & q_2q_5 & q_4q_5 \\ q_1q_2 & q_1q_4 & q_2q_4 \\ q_3q_4 & q_3q_5 & q_4q_5 \\ q_2q_3 & q_2q_4 & q_3q_4 \\ q_1q_3 & q_1q_6 & q_3q_6 \\ q_1q_3 & q_1q_7 & q_3q_7 \\ q_2q_4 & q_2q_7 & q_4q_7 \\ q_1q_5 & q_1q_7 & q_5q_7 \\ q_3q_5 & q_3q_7 & q_5q_7 \\ q_1q_3 & q_1q_8 & q_3q_8 \\ q_2q_4 & q_2q_8 & q_4q_8 \\ q_1q_5 & q_1q_8 & q_5q_8 \\ q_3q_5 & q_3q_8 & q_5q_8 \\ q_1q_3 & q_1q_9 & q_3q_9 \\ q_2q_4 & q_2q_9 & q_4q_9 \\ q_1q_5 & q_1q_9 & q_5q_9 \\ q_3q_5 & q_3q_9 & q_5q_9 \\ q_1q_3 & q_1q_{10} & q_3q_{10} \\ q_2q_4 & q_2q_{10} & q_4q_{10} \\ q_1q_5 & q_1q_{10} & q_5q_{10} \\ q_3q_5 & q_3q_{10} & q_5q_{10} \end{pmatrix} \tag{3.50}$$

Eq. 3.50 shows that there are 30, 3-local terms in Eq. 3.49 and three different ways to collapse each of those 3-local terms. In general, the problem of choosing the most efficient 2-local terms to collapse this function is NP-Complete. This becomes evident if we represent our problem as an element cover on a bipartite graph. Suppose we relabel each 3-local term

on the left as "set" 1-30, denoted as $S_1 S_2 ... S_{30}$. We can then make the following bipartite graph which connects the 3-local terms to the 2-local terms which collapse them.

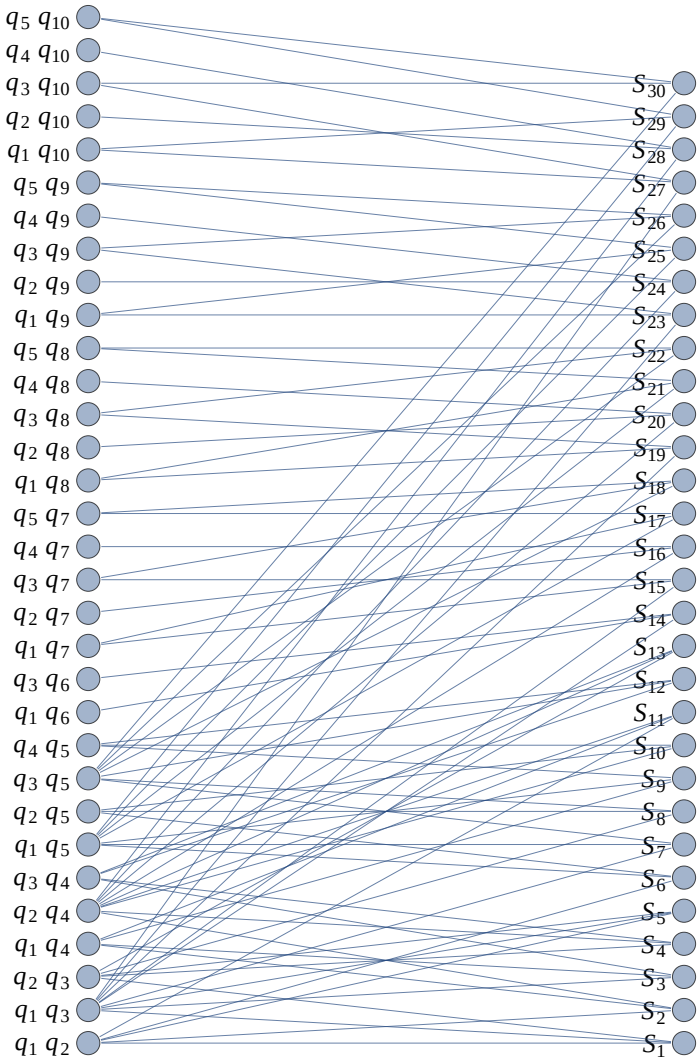



Figure 3.17: A bipartite graph connecting the 3-local terms ($S_n$) in Eq. 3.49 to the 2-local terms ($q_i q_j$) which collapse them.

Fig. 3.19 shows that we can now restate the problem in the following way: "choose the fewest number of 2-local terms (on the left) which covers all 3-local terms (on the right)

with at least one edge." In general, this problem is isomorphic to the canonical "hitting set" problem which is equivalent to set cover, one of Karp's 21 NP-Complete Problems [66, 76, 178]. However, we have specifically kept this issue in mind when creating the turn-ancilla representation in such a way as to guarantee that it is easy to find a relatively efficient solution to this problem. Accordingly, our experience has been that a greedy local-search algorithm performs very well.

The explanation for this is simple: each 3 or 4-local term will contain no more than 1 ancillary bit; thus, to cover all 3 and 4-local terms we can focus entirely on the physical bits (in this case, bits 1-5). In alternative mappings not presented here we have frequently encountered extremely difficult instances of the hitting set problem during the reduction process. In these situations one should see [250] for a very efficient algorithm which can exactly solve hitting cover in $O\left(1.23801^n\right)$.

## 3.7 Example Encoding

A primary goal of this review is to elucidate an efficient process for encoding chemical physics problems into a form suitable for quantum computation. In addition to providing the alternatives for the solution of the lattice heteropolymer problem in quantum devices, we seek to provide a general explanation of considerations for constructing energy functions for these devices. These have many possible applications for solving problems related to statistical mechanics on the device. In this section, we will complete our review by demonstrating the final steps required to embed a small instance of a particular lattice protein problem into a QUBO Hamiltonian.

The Hamiltonians and the number of resources presented here correspond to the minimum amount of resources needed assuming the device can handle many-body interactions

as is the case for NMR quantum computers or trapped ions. The hierarchical experimental proposals presented here work for lattice folding under no external constraints, i.e., amino acid chains in "free space" [1]. As a final step we will reduce these Hamiltonians to a 2-local form specifically design for the Dwave One used in [221, 211, 86, 154]. The final Hamiltonian we present is more efficient than that used in [221] as we have since realized several tricks to make the energy function more compact.

### 3.7.1 Previous experimental implementation

Throughout this review we have referred to an experimental implementation of quantum annealing to solve lattice heteropolymer problems in [221]. The quantum hardware employed consists of 16 units of a recently characterized eight qubit unit cell [154, 135]. Post-fabrication characterization determined that only 115 qubits out of the 128 qubit array can be reliably used for computation. The array of coupled superconducting flux qubits is, effectively, an artificial Ising spin system with programmable spin-spin couplings and transverse magnetic fields. It is designed to solve instances of the following (NP-hard) classical optimization problem: given a set of local longitudinal fields ($h_i$) and an interaction matrix ($J_{ij}$), find the assignment $\mathbf{s} = \mathbf{s_1 s_2 s_3 ... s_N}$ , that minimizes the objective function $E(\mathbf{s})$, where,

$$E\left(\mathbf{s}\right) = \sum_{1 \leq i \leq N} h_i s_i + \sum_{1 \leq i \leq j \leq N} J_{ij} s_i s_j \tag{3.51}$$

and $s_i \in -1, 1$. Thus, the solution to this problem, $\mathbf{s}$, can be encoded into the ground-state wavefunction of the quantum Hamiltonian,

$$\mathcal{H}_p = \sum_{1 \leq i \leq N} h_i \sigma_i^z + \sum_{1 \leq i \leq j \leq N} J_{ij} \sigma_i^z \sigma_j^z. \tag{3.52}$$

[1] External interactions could also be included as presented and verified experimentally in [221].

Quantum annealing exploits the adiabatic theorem of quantum mechanics, which states that a quantum system initialized in the ground state of a time-dependent Hamiltonian remains in the instantaneous ground state, as long as it is driven sufficiently slowly. Since the ground state of $\mathcal{H}_p$ encodes the solution to the optimization problem, the idea behind quantum annealing is to adiabatically prepare this ground state by initializing the quantum system in some easy-to-prepare ground state, $\mathcal{H}_b$. In this case, $\mathcal{H}_b$ corresponds to a superposition of all states of the computational basis. The system is driven slowly to the problem Hamiltonian, $\mathcal{H}(\tau = 1) \approx \mathcal{H}_p$. Deviations from the ground-state are expected due to deviations from adiabaticity, as well as thermal noise and imperfections in the implementation of the Hamiltonian.

Using the encoding methods discussed here, the authors were able to encode and to solve the global minima solution for small tetrapeptide and hexapeptide chains under several experimental schemes involving 5 and 8 qubits for four-amino-acid sequence (Hydrophobic-Polar model) and 5, 27, 28, and 81 qubits experiments for the six-amino-acid sequence under the Miyazawa-Jernigan model for general pairwise interactions.

### 3.7.2 Six unit Miyazawa-Jernigan protein

The example we will present here is a different encoding of the largest problem performed in [221]: the Miyazawa-Jernigan (MJ) protein, Proline-Serine-Valine-Lysine-Methionine-Alanine (PSVKMA) on a $2D$ lattice. We will use the pair-wise nearest-neighbor MJ interaction energies presented in Table 3 of [199] and shown in Fig. 3.18.

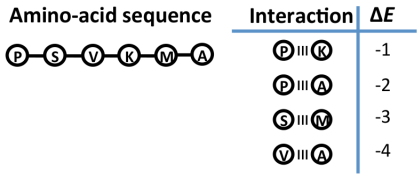



Figure 3.18: Interaction matrix for our protein in the MJ model.

We will use the turn ancilla construction for our energy function and constrain the first three virtual bits to 010, as before. Recall that the turn ancilla construction requires $2N-5$ physical information bits; thus, our 6-unit MJ protein will be encoded into 7 bits.

### $E_{back}(\boldsymbol{q})$ for 6-unit SAW on $2D$ lattice

Using Eq. 3.13, we find that our 6-unit protein has the backwards energy function,

$$\begin{aligned} E_{back}(\boldsymbol{q}) = \lambda_{back}(&q_1q_2 - 2q_1q_3q_2 + 2q_3q_2 - 2q_3q_4q_2 - 2q_3q_5q_2 \\ &+ 4q_3q_4q_5q_2 - 2q_4q_5q_2 + q_5q_2 + q_3q_4 - 2q_3q_4q_5 + 2q_4q_5 - 2q_4q_5q_6 \\ &+ q_5q_6 + q_4q_7 - 2q_4q_5q_7 - 2q_4q_6q_7 + 4q_4q_5q_6q_7 - 2q_5q_6q_7 + q_6q_7). \end{aligned} \tag{3.53}$$

Soon, we will discuss how to choose the appropriate value for $\lambda_{back}$ but for now we simply note that $\lambda_{back}$ and $\lambda_{overlap}$ penalize the same illegal folds; thus we realize that $\lambda_{back} = \lambda_{overlap}$.

### $E_{overlap}(\boldsymbol{q})$ for 6-unit SAW on $2D$ lattice

Using Eq. 3.30, we calculate the overlap energy function as,

$$
\begin{aligned}
E_{overlap}(\boldsymbol{q}) = \lambda_{overlap}(&96q_2q_1 - 96q_2q_3q_1 - 64q_3q_1 - 64q_2q_4q_1 + 64q_2q_3q_4q_1 - 96q_3q_4q_1 + 96q_4q_1 \\
&- 96q_2q_5q_1 + 64q_2q_4q_5q_1 - 96q_4q_5q_1 - 64q_5q_1 - 48q_2q_6q_1 + 32q_2q_3q_6q_1 - 48q_3q_6q_1 + 32q_3q_4q_6q_1 \\
&- 48q_4q_6q_1 + 32q_2q_5q_6q_1 + 32q_4q_5q_6q_1 - 48q_5q_6q_1 + 72q_6q_1 - 48q_2q_7q_1 - 48q_3q_7q_1 + 32q_2q_4q_7q_1 \\
&- 48q_4q_7q_1 + 96q_3q_5q_7q_1 - 48q_5q_7q_1 + 32q_2q_6q_7q_1 + 32q_4q_6q_7q_1 - 48q_6q_7q_1 - 8q_7q_1 - 8q_3q_{10} \\
&+ 64q_3q_8q_1 + 64q_5q_8q_1 - 32q_8q_1 + 32q_3q_9q_1 + 32q_5q_9q_1 - 16q_9q_1 + 16q_3q_{10}q_1 + 16q_5q_{10}q_1 - 8q_{10}q_1 \\
&+ 8q_3q_{11}q_1 + 8q_5q_{11}q_1 - 4q_{11}q_1 + 64q_3q_{12}q_1 + 64q_5q_{12}q_1 + 64q_7q_{12}q_1 - 96q_{12}q_1 + 32q_3q_{13}q_1 \\
&+ 32q_5q_{13}q_1 + 32q_7q_{13}q_1 - 48q_{13}q_1 + 16q_3q_{14}q_1 + 16q_5q_{14}q_1 + 16q_7q_{14}q_1 - 24q_{14}q_1 + 8q_3q_{15}q_1 \\
&+ 8q_5q_{15}q_1 + 8q_7q_{15}q_1 - 12q_{15}q_1 + 64q_1 + 144q_2 + 96q_2q_3 + 64q_3 - 64q_2q_4 - 64q_2q_3q_4 + 96q_3q_4 + 144q_4 \\
&+ 96q_2q_5 - 96q_2q_3q_5 - 64q_3q_5 - 64q_2q_4q_5 + 64q_2q_3q_4q_5 - 96q_3q_4q_5 + 96q_4q_5 + 64q_5 - 8q_2q_6 - 48q_2q_3q_6 \\
&+ 72q_3q_6 - 48q_2q_4q_6 - 48q_3q_4q_6 - 8q_4q_6 - 48q_2q_5q_6 + 32q_2q_3q_5q_6 - 48q_3q_5q_6 + 32q_3q_4q_5q_6 - 48q_4q_5q_6 \\
&+ 72q_5q_6 + 36q_6 + 72q_2q_7 - 48q_2q_3q_7 - 8q_3q_7 - 48q_2q_4q_7 + 32q_2q_3q_4q_7 - 48q_3q_4q_7 + 72q_4q_7 - 48q_2q_5q_7 \\
&- 48q_3q_5q_7 + 32q_2q_4q_5q_7 - 48q_4q_5q_7 - 8q_5q_7 - 48q_2q_6q_7 + 32q_2q_3q_6q_7 - 48q_3q_6q_7 + 32q_3q_4q_6q_7 \\
&+ 32q_2q_5q_6q_7 + 32q_4q_5q_6q_7 - 48q_5q_6q_7 + 72q_6q_7 + 36q_7 - 96q_2q_8 - 32q_3q_8 + 64q_2q_4q_8 - 96q_4q_8 \\
&- 32q_5q_8 - 32q_8 - 48q_2q_9 - 16q_3q_9 + 32q_2q_4q_9 - 48q_4q_9 + 32q_3q_5q_9 - 16q_5q_9 + 64q_8q_9 - 32q_9 - 24q_2q_{10} \\
&+ 16q_2q_4q_{10} - 24q_4q_{10} + 16q_3q_5q_{10} - 8q_5q_{10} + 32q_8q_{10} + 16q_9q_{10} - 20q_{10} - 12q_2q_{11} - 4q_3q_{11} + 8q_2q_4q_{11} \\
&- 12q_4q_{11} + 8q_3q_5q_{11} - 4q_5q_{11} + 16q_8q_{11} + 8q_9q_{11} + 4q_{10}q_{11} - 11q_{11} - 96q_2q_{12} - 96q_3q_{12} + 64q_2q_4q_{12} \\
&- 96q_4q_{12} + 64q_3q_5q_{12} - 96q_5q_{12} + 64q_2q_6q_{12} + 64q_4q_6q_{12} - 96q_6q_{12} + 64q_3q_7q_{12} + 64q_5q_7q_{12} \\
&- 96q_7q_{12} + 64q_{12} - 48q_2q_{13} - 48q_3q_{13} + 32q_2q_4q_{13} - 48q_4q_{13} + 32q_3q_5q_{13} - 48q_5q_{13} + 32q_2q_6q_{13} \\
&+ 32q_4q_6q_{13} - 48q_6q_{13} + 32q_3q_7q_{13} + 32q_5q_7q_{13} - 48q_7q_{13} + 64q_{12}q_{13} + 16q_{13} - 24q_2q_{14} - 24q_3q_{14} \\
&+ 16q_2q_4q_{14} - 24q_4q_{14} + 16q_3q_5q_{14} - 24q_5q_{14} + 16q_2q_6q_{14} + 16q_4q_6q_{14} - 24q_6q_{14} + 16q_3q_7q_{14} \\
&+ 16q_5q_7q_{14} - 24q_7q_{14} + 32q_{12}q_{14} + 16q_{13}q_{14} + 4q_{14} - 12q_2q_{15} - 12q_3q_{15} + 8q_2q_4q_{15} - 12q_4q_{15} \\
&+ 8q_3q_5q_{15} - 12q_5q_{15} + 8q_2q_6q_{15} + 8q_4q_6q_{15} - 12q_6q_{15} + 8q_3q_7q_{15} + 8q_5q_7q_{15} - 12q_7q_{15} \\
&+ 16q_{12}q_{15} + 8q_{13}q_{15} + 4q_{14}q_{15} + q_{15} - 48q_4q_6q_7 + 64q_3q_5q_8). \qquad (3.54)
\end{aligned}
$$

We notice that as discussed in Sec. 3.6, all the 3-local terms here contain at least two physical information qubits (i.e. $q_1$ through $q_7$).

### $E_{pair}(\boldsymbol{q})$ for MJ-model PSVKMA

Using the $J$ matrix as defined in Eq. 3.32 we calculate the pair-wise energy function as,

$$\begin{aligned}
E_{pair}(\boldsymbol{q}) = &-4q_2q_{16} + 4q_1q_3q_{16} + 3q_{16} - 8q_1q_{17} - 16q_2q_{17} + 8q_1q_3q_{17} - 8q_3q_{17} \\
&+ 8q_2q_4q_{17} - 16q_4q_{17} + 8q_1q_5q_{17} + 8q_3q_5q_{17} - 8q_5q_{17} + 8q_2q_6q_{17} + 8q_4q_6q_{17} - 16q_6q_{17} \\
&+ 8q_1q_7q_{17} + 8q_3q_7q_{17} + 8q_5q_7q_{17} - 8q_7q_{17} + 30q_{17} - 12q_1q_{18} - 12q_2q_{18} + 12q_1q_3q_{18} \\
&- 12q_3q_{18} + 12q_2q_4q_{18} - 12q_4q_{18} + 12q_1q_5q_{18} + 12q_3q_5q_{18} - 12q_5q_{18} + 21q_{18} - 16q_2q_{19} \\
&- 16q_3q_{19} + 16q_2q_4q_{19} - 16q_4q_{19} + 16q_3q_5q_{19} - 16q_5q_{19} + 16q_2q_6q_{19} + 16q_4q_6q_{19} \\
&- 16q_6q_{19} + 16q_3q_7q_{19} + 16q_5q_7q_{19} - 16q_7q_{19} + 28q_{19}. \qquad (3.55)
\end{aligned}$$

### Setting $\lambda$ penalty values

Finally, we will discuss how one chooses the correct penalty values for the energy function. This is a crucial step if one wishes to implement the algorithm experimentally as all currently available architectures for adiabatic quantum annealing have limited coupler resolution. That is, quantum annealing machines cannot realize arbitrary constant values for the QUBO expression. Thus, it is very important that one chooses the lowest possible penalty values which still impose the correct constraints. In our problem we choose the value of $\lambda_{overlap}$ by asking ourselves: what is the greatest possible amount that any overlap could *lower* the system energy? In general, a very conservative upper bound can be obtained by simply summing together every $J$ matrix element (which would mean that a single overlap allowed every single possible interaction to occur); in our problem this upper-bound would be -10. Thus, we can set $\lambda_{overlap} = +10$.

**Reduction to 2-local**

Using a standard greedy search algorithm we find that an efficient way to collapse this energy function to 2-local is to make ancilla with the qubit pairs,

$$\begin{aligned} q_2 q_4 &\rightarrow q_{20} \\ q_1 q_3 &\rightarrow q_{21} \\ q_3 q_5 &\rightarrow q_{22} \\ q_1 q_5 &\rightarrow q_{23} \\ q_2 q_6 &\rightarrow q_{24} \\ q_4 q_6 &\rightarrow q_{25} \\ q_3 q_7 &\rightarrow q_{26} \\ q_5 q_7 &\rightarrow q_{27} \\ q_1 q_7 &\rightarrow q_{28} \quad . \end{aligned} \tag{3.56}$$

There is one issue left to discuss - the value of $\delta_n$ in Eq. 3.48. The purpose of $\delta_n$ is to constrain the reductions in Eq. 3.56 so that the value of the ancillary bit actually corresponds to the product of the two bits it is supposed to represent. In order for Eq. 3.48 to work we must choose $\delta_n$ which is large enough so that a violation of the reduction we desire will always raise the system energy. Thus, we must ensure that $\delta_n$ is large enough so that configurations which do not conform to the reduction are penalized by an amount higher than the largest penalty they could avoid and larger in magnitude than the largest energy reduction they could achieve with the illegal move. Of course, finding the exact minimum value of $E(\boldsymbol{q})$ is as difficult as minimizing $E(\boldsymbol{q})$ (our goal). Instead, we can simply make an upper-bound for the penalty by setting it equal to one plus either the sum of the absolute value of all psuedo-boolean coefficients corresponding to the variables being collapsed in $E(\boldsymbol{q})$ (whichever sum is larger).

### QUBO Matrix and Solutions

After reduction of the energy function to 2-local, we arrive at the final pseudo-boolean energy function. Instead of writing out the entire pseudo-boolean expression we will instead provide a matrix containing all of the coefficients of 1-local terms on the diagonal and 2-local terms in the upper triangular portion of this matrix. This representation is known as the QUBO matrix and contains all of the couplings needed for experimental implementation and is shown in Eq. 3.57. Note that the full pseudo-boolean expression contains one constant term that we drop in the matrix representation. This constant has a value of $C = 180$ for this particular problem.

$$
\begin{pmatrix}
320 & 485 & 42962 & 480 & 42962 & 360 & 42962 & -160 & -80 & -40 & -20 & -480 & -240 & -120 & -60 & 0 & -8 & -12 & 0 & -320 & -85924 & 0 & -85924 & -240 & -240 & 0 & 0 & -85924 \\
0 & 720 & 490 & 42962 & 485 & 42962 & 360 & -480 & -240 & -120 & -60 & -480 & -240 & -120 & -60 & -4 & -16 & -12 & -16 & -85924 & -490 & -490 & -480 & -85924 & 0 & -240 & -240 & -240 \\
0 & 0 & 320 & 485 & 42962 & 360 & 42962 & -160 & -80 & -40 & -20 & -480 & -240 & -120 & -60 & 0 & -8 & -12 & -16 & -330 & -85924 & -85924 & 0 & -240 & -240 & -85924 & 0 & 0 \\
0 & 0 & 0 & 720 & 490 & 42962 & 365 & -480 & -240 & -120 & -60 & -480 & -240 & -120 & -60 & 0 & -16 & -12 & -16 & -85924 & -480 & -490 & -480 & 0 & -85924 & -240 & -250 & -240 \\
0 & 0 & 0 & 0 & 320 & 365 & 42962 & -160 & -80 & -40 & -20 & -480 & -240 & -120 & -60 & 0 & -8 & -12 & -16 & -330 & 0 & -85924 & -85924 & -240 & -250 & 0 & -85924 & 0 \\
0 & 0 & 0 & 0 & 0 & 180 & 365 & 0 & 0 & 0 & 0 & -480 & -240 & -120 & -60 & 0 & -16 & 0 & -16 & -240 & -240 & -240 & -240 & -85924 & -85924 & -240 & -250 & -240 \\
0 & 0 & 0 & 0 & 0 & 0 & 180 & 0 & 0 & 0 & 0 & -480 & -240 & -120 & -60 & 0 & -8 & 0 & -16 & -240 & -240 & -240 & -240 & -240 & -250 & -85924 & -85924 & -85924 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & -160 & 320 & 160 & 80 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 320 & 320 & 320 & 320 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -160 & 80 & 40 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 160 & 160 & 160 & 160 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -100 & 20 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 80 & 80 & 80 & 80 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -55 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 40 & 40 & 40 & 40 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 320 & 320 & 160 & 80 & 0 & 0 & 0 & 0 & 320 & 320 & 320 & 320 & 320 & 320 & 320 & 320 & 320 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 80 & 80 & 40 & 0 & 0 & 0 & 0 & 160 & 160 & 160 & 160 & 160 & 160 & 160 & 160 & 160 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 20 & 20 & 0 & 0 & 0 & 0 & 80 & 80 & 80 & 80 & 80 & 80 & 80 & 80 & 80 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 & 0 & 0 & 0 & 40 & 40 & 40 & 40 & 40 & 40 & 40 & 40 & 40 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 & 0 & 0 & 0 & 0 & 4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 30 & 0 & 0 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 21 & 0 & 12 & 12 & 12 & 12 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 28 & 16 & 0 & 16 & 0 & 16 & 16 & 16 & 16 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128566 & 320 & 340 & 320 & 0 & 0 & 160 & 160 & 160 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128566 & 0 & 0 & 160 & 160 & 0 & 480 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128566 & 0 & 160 & 160 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128566 & 160 & 160 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128846 & 0 & 160 & 160 & 160 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128846 & 160 & 180 & 160 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128846 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128846 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 128846
\end{pmatrix}
\tag{3.57}
$$

Taking the matrix in Eq. 3.57 as $Q$, we can write the total energy of a given solution (denoted by $\boldsymbol{q}$) as,

$$E(\boldsymbol{q}) = \boldsymbol{q}Q\boldsymbol{q}. \tag{3.58}$$

The problem is now ready for its implementation on a quantum device. For our particular problem instance the solution string is given by the bit string,

$$0,0,0,1,0,1,1,1,1,0,0,1,1,1,0,0,1,0,1,0,0,0,0,0,1,0,0,0. \tag{3.59}$$

The energy given by Eq. 3.58 is $-186$. In the original expression this corresponds to an energy of $C - 186 = 180 - 186 = -6$. Let's confirm that this is accurate to the MJ model. Looking only at the physical information bits and prepending the first three constant bits (010) we see that the bit string prescribes the following fold:

$$\boldsymbol{q} = \underbrace{01}_{\text{right}}\underbrace{00}_{\text{down}}\underbrace{00}_{\text{down}}\underbrace{10}_{\text{left}}\underbrace{11}_{\text{up}} \tag{3.60}$$

which corresponds to the fold,

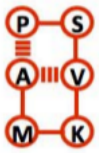



Figure 3.19: The solution to our example problem for MJ protein PSVKMA.

## 3.8 Conclusion

As both traditional and quantum computer science continue to advance as fields, domain scientists from all disciplines need to develop new ways of representing problems in order to leverage state-of-the-art computational tools. In this review, we discussed strategies and techniques for solving lattice heteropolymer problems with some of these tools.

While the lattice heteropolymer model is widely applicable to many problems, the general principles used to optimally encode and constrain this particular application are fairly universal for discrete optimization problems in the physical sciences.

We focused on three mappings: "turn ancilla", "turn circuit" and "diamond". The turn ancilla mapping is the best mapping in terms of the scaling of the number of resources for large instances, thus making it ideal for benchmark studies of lattice folding using (heuristic) solvers for pseudo-boolean minimization. Additionally, this method shows how one can use ancilla variables to construct a fitness function with relatively few constraints per clause (i.e. low-locality). With ancilla variables even an extremely simple encoding, such as the turn encoding, can be used to construct a complicated energy function. While some of the particular tricks employed to optimize the efficiency of this mapping, such as introducing the backwards penalty, are specific to lattice heteropolymers, the general logic behind these tricks is much more universal.

The turn circuit mapping is the most compact of all three mappings. The extremely efficient use of variables (qubits) makes it ideal for benchmark experiments on quantum devices which can handle many body couplings. Moreover, the turn circuit method demonstrates how one can construct an elaborate energy function by utilizing logic circuits to put together a high-local fitness function of arbitrary complexity without ancilla variables. While different problems may involve different circuits, the underlying strategy is very broadly applicable.

The diamond encoding illustrates a strategy for producing an extremely under-constrained optimization problem. Furthermore, this method demonstrates that even fairly complex energy functions can be represented as natively 2-local functions if one is willing to sacrifice efficiency. Many quantum devices can only couple bits pairwise; thus, this is a very impor-

tant quality of the diamond encoding. Finally, if one uses another, more efficient encoding, we explain how reductions can be used to replace high-local terms with 2-local terms in an optimally efficient fashion but at the cost of needing very high coupler resolution. The relatively few constraints in the diamond encoding make it a natural choice for exact or heuristic ILP and W-SAT solvers.

These three strategies elucidate many of the concepts that we find important when producing problems suitable for the D-Wave device utilized in [221]. Accordingly, as quantum information science continues to develop, we hope that the methods discussed in this review will be useful to scientists wishing to leverage similar technology for the solution of discrete optimization problems.

## Acknowledgements

The authors thank Joseph Goodknight for help editing this chapter. This research was sponsored by United States Department of Defense, grant M1144-201167-DS. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the U.S. Government.

## Chapter 4

# Bayesian Network Structure Learning Using Quantum Annealing

Apart from minor modifications, this chapter originally appeared as [215]:

"Bayesian Network Structure Learning Using Quantum Annealing". Bryan O'Gorman, Alejandro Perdomo-Ortiz, Ryan Babbush, Alán Aspuru-Guzik and Vadim Smelyanskiy. *European Physical Journal Special Topics*. Volume 225, Number 1: 163-188. 2015.

## Abstract

We introduce a method for the problem of learning the structure of a Bayesian network using the quantum adiabatic algorithm. We do so by introducing an efficient reformulation of a standard posterior-probability scoring function on graphs as a pseudo-Boolean function, which is equivalent to a system of 2-body Ising spins, as well as suitable penalty terms for enforcing the constraints necessary for the reformulation; our proposed method requires $\mathcal{O}(n^2)$ qubits for $n$ Bayesian network variables. Furthermore, we prove lower bounds on the necessary weighting of these penalty terms. The logical structure resulting from the mapping has the appealing property that it is instance-independent for a given number of Bayesian network variables, as well as being independent of the number of data cases.

## 4.1 Introduction

Bayesian networks are a widely used probabilistic graphical model in machine learning [171]. A Bayesian network's structure encapsulates conditional independence within a set of random variables, and, equivalently, enables a concise, factored representation of their joint probability distribution. There are two broad classes of computational problems associated with Bayesian networks: inference problems, in which the goal is to calculate a probability distribution or the mode thereof given a Bayesian network and the state of some subset of the variables; and learning problems, in which the goal is to find the Bayesian network most likely to have produced a given set of data. Here, we focus on the latter, specifically the problem of Bayesian network structure learning. Bayesian network structure learning has been applied in fields as diverse as the short-term prediction of solar-flares [294] and the discovery of gene regulatory networks [93, 111]. The problem of learning the most likely structure to have produced a given data set, with reasonable formal assumptions to be enumerated later, is known to be NP-complete [68], so its solution in practice requires the use of heuristics.

Quantum annealing is one such heuristic. Though efficient quantum algorithms for certain problems are exponentially faster than their classical counterpart, it is believed that quantum computers cannot efficiently solve NP-complete problems [208]. However, there exist quantum algorithms that have a provable speedup over classical ones [124, 256]. There is therefore reason to believe quantum-mechanical effects such as tunneling could provide a polynomial speedup over classical computation for some sets of problems. The recent availability of quantum annealing devices from D-Wave Systems has sparked interest in the experimental determination of whether or not the current generation of the device provides such speedup [43, 44, 233, 273]. While there exists prior work related to "quantum Bayesian

networks"[267] and the "quantum computerization" of classical Bayesian network methods [268], the results presented here are unrelated.

In this paper, we describe how to efficiently map a certain formulation of Bayesian Network Structure Learning (BNSL) to Quadratic Unconstrained Binary Optimization (QUBO). The QUBO formalism is useful because it is mathematically equivalent to that of a set Ising spins with arbitrary 2-body interactions, which can be mapped to the Ising spins with a limited 2-body interaction graph as implementable by physical quantum annealing devices. Similar mappings have been developed and implemented for lattice protein folding [18, 221], planning and scheduling [229], fault diagnosis [222], graph isomorphism [112], training a binary classifier [13, 209], and the computation of Ramsey numbers [39].

To map BNSL to QUBO, we first encode all digraphs using a set of Boolean variables, each of which indicates the presence or absence of an arc (i.e. directed edge), and define a pseudo-Boolean function on those variables that yields the score of the digraph encoded therein so long as it satisfies the necessary constraints. This function is not necessarily quadratic, and so we apply standard methods to quadratize (i.e. reduce the degree to two) using ancillary variables. We then introduce ancillary variables and add additional terms to the pseudo-Boolean function corresponding to constraints, each of which is zero when the corresponding constraint is satisfied and positive when it is not. The resulting QUBO instance is defined over $\mathcal{O}(n^2)$ Boolean variables when mapped from a BNSL instance with $n$ Bayesian network variables. Interestingly, the structure of the QUBO is instance-independent for a fixed BNSL size. Because embedding the structure of QUBO into physical hardware is generally computationally difficult, this is an especially appealing feature of the mapping.

We also show sufficient lower bounds on penalty weights used to scale the terms in the Hamiltonian that penalize invalid states, like those containing a directed cycle or with parent sets larger than allowed. In a physical device, setting the penalty weights too high is counterproductive because there is a fixed maximum energy scale. The stronger the penalty weights, the more the logical energy spectrum is compressed, which is problematic for two reasons: first, the minimum gap, with which the running time of the algorithm scales inversely, is proportionally compressed, and, second, the inherently limited precision of a physical device's implementation of the interaction strengths prevents sufficient resolution of logical states close in energy as the spectrum is compressed.

The utility of the mapping from BNSL to QUBO introduced here is not limited to quantum annealing. Indeed, the methods used here were motivated by a previous mapping of the same problem to weighted MAX-SAT [83]. Existing simulated annealing code is highly optimized [150] and may be applied to QUBO instances derived from our mapping. In that case, there is no need to quadratize, because simulated annealing does not have the limitation to 2-body interactions that physical devices do. With respect to penalty weights, while simulated annealing does not have the same gap and precision issues present in quantum annealing, there may still be reason to avoid setting the penalty weights too high. Because the bits corresponding to arcs with different i.e. terminal vertices do not interact directly, many valid states are separated by invalid ones, and so penalty weights that are too strong may erect barriers that tend to produce basins of local optima. While simulated annealing directly on digraph structures is possible, mapping to QUBO and performing simulated annealing in that form has the advantage that it enables the exploitation of existing, highly optimized code, as well as providing an alternative topology of the solution space and energy landscape.

BNSL has a special property that makes it especially well-suited for the application of heuristics such as QA: Unlike in other problems where anything but the global minimum is undesirable or those in which an approximate solution is sufficient, in BNSL there is utility in having a set of high scoring DAGs. The scoring function encodes the posterior probability, and so sub- but near-optimal solution may be almost as probable as the global optimum. In practice, because quantum annealing is an inherently stochastic procedure, it is run many times for the same instance, producing a set of low-energy states. In cases where the BN structure is learned for the purpose of doing inference on it, a high-scoring subset of many quantum annealing runs can utilized by performing Bayesian model averaging, in which inference is done on the set of likely BNs and the results averaged proportionally.

In Section 4.2, we review the formalism of Bayesian networks and BNSL (4.2.1) and quantum annealing (4.2.2), elucidating the features that make the latter suitable for finding solutions of the former. In Section 4.3, we develop an efficient and instance-independent mapping from BNSL to QUBO. In Section 4.4, we provide sufficient lower bounds on the penalty weights in the aforementioned mapping. In Section 4.5, we discuss useful features of the mapping and conclude. In the Appendix, we prove the sufficiency of the lower bounds given; the methods used to do so may be useful in mappings for other problems.

## 4.2 Background

### 4.2.1 Bayesian Network Structure Learning

A Bayesian network (BN) is a probabilistic graphical model for a set of random variables that encodes their joint probability distribution in a more compact way and with fewer parameters than would be required otherwise by taking into account conditional independences among the variables. It consists of both a directed acyclic graph (DAG) whose

vertices correspond to the random variables and an associated set of conditional probabilities for each vertex. Here and throughout the literature, the same notation is used for both a random variable and its corresponding vertex, and the referent will be clear from context.

Formally, a BN $B$ for $n$ random variables $\mathbf{X} = (X_i)_{i=1}^n$ is a pair $(B_S, B_P)$, where $B_S$ is a DAG representing the structure of the network and $B_P$ is the set of conditional probabilities $\{p(X_i|\Pi_i(B_S))|1 \leq i \leq n\}$ that give the probability distribution for the state of a variable $X_i$ conditioned on the joint state of its parent set $\Pi_i(B_S)$ (those variables for which there are arcs in the structure $B_S$ from the corresponding vertices to that corresponding to $X_i$; we will write simply $\Pi_i$ where the structure is clear from context). Let $r_i$ denote the number of states of the variable $X_i$ and $q_i = \prod_{j \in \Pi_i} r_j$ denote the number of joint states of the parent set $\Pi_i$ of $X_i$ (in $B_S$). Lowercase variables indicate realizations of the corresponding random variable; $x_{ik}$ indicates the $k$-th state of variable $X_i$ and $\pi_{ij}$ indicates the $j$-th joint state of the parent set $\Pi_i$. The set of conditional probabilities $B_P$ consists of $n$ probability distributions $\left((\theta_{ij})_{j=1}^{q_i}\right)_{i=1}^n$, where $\theta_{ij} = (\theta_{ijk})_{k=1}^{r_i}$ is the conditional probability distribution for the states $(x_{ik})_{k=1}^{r_i}$ of the variable $X_i$ given the joint state $\pi_{ij}$ of its parents $\Pi_i$ (i.e. $p(x_{ik}|\pi_{ij}) = \theta_{ijk}$).

Given a database $D = \{\mathbf{x}_i|1 \leq i \leq N\}$ consisting of $N$ cases, where each $\mathbf{x}_i$ denotes the state of all variables $\mathbf{X}$, the goal is to find the structure that maximizes the posterior distribution $p(B_S|D)$ out of all possible structures. By Bayes's Theorem,

$$p(B_S|D) = \frac{p(D|B_S)p(B_S)}{p(D)}. \tag{4.1}$$

The marginal probability of the database $p(D)$ is the same for all structures, so assuming that each structure is equally likely, this simplifies to

$$p(B_S|D) \propto p(D|B_S). \tag{4.2}$$

In Section 4.3.5, we describe how to account for certain types of non-uniform prior distributions over the graph structures. With certain further reasonable assumptions, namely multinomial sampling, parameter independence and modularity, and Dirichlet priors, the latter conditional probability is

$$p(D|B_S) = \prod_{i=1}^{n} \prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(N_{ij} + \alpha_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(N_{ijk} + \alpha_{ijk})}{\Gamma(\alpha_{ijk})}, \tag{4.3}$$

where $N_{ijk}$ is the number of cases in $D$ such that variable $X_i$ is in its $k$-th state and its parent set $\Pi_i$ is in its $j$-th state, $N_{ij} = \sum_{k=1}^{r_i} N_{ijk}$, $\alpha_{ijk}$ is the hyperparameter for $\theta_{ijk}$ in the Dirichlet distribution from which $\theta_{ij}$ is assumed to be drawn, and $\alpha_{ij} = \sum_{k=1}^{r_i} \alpha_{ijk}$ [143].

Given a database $D$, our goal is equivalent to that of finding the structure with the largest likelihood, i.e. the structure that yields the largest probability of the given database conditioned on that structure. We do this by encoding all structures into a set of bits and defining a quadratic pseudo-Boolean function on those bits and additional ancillary bits whose minimizing bitstring encodes the structure with the largest posterior probability.

### 4.2.2 Quantum Annealing

Quantum annealing is a method for finding the minimum value of a given objective function. It is the quantum analogue of classical simulated annealing, where the computation is driven by quantum, rather than thermal, fluctuations [159]. A similar procedure is called adiabatic quantum computation, in which the adiabatic interpolation of a Hamiltonian whose ground state is easily prepared to one whose ground state encodes the solution to the desired optimization problem guarantees that final state is indeed the ground state of the latter [104]. The formalism for both is similar, and the methods described here are useful for both. Specifically, the time-dependent Hamiltonian is

$$H(t) = A(t)H_0 + B(t)H_1, \tag{4.4}$$

for $0 \leq t \leq T$, where $H_0$ is the initial Hamiltonian, $H_1$ is the final Hamiltonian, $A(t)$ is a real monotonic function such that $A(0) = 1$ and $A(T) = 0$, and $B(t)$ is a real monotonic function such that $B(0) = 0$ and $B(T) = 1$. The adiabatic theorem states that if the system starts in the ground state of $H_0$ and $H(t)$ varies slowly enough, then the system will be in the ground state of $H_1$ at time $T$. Using this procedure to solve an optimization problem entails the construction of $H_1$ such that its ground state encodes the optimal solution. In practice, arbitrary Hamiltonians are difficult to implement, but this is ameliorated by results showing the ability to effectively implement arbitrary Hamiltonians using physically-realizable connectivity through various gadgetry with reasonable overhead [216, 161].

The main contribution of this paper is a construction of $H_1$ such that its ground state encodes the solution for a given instance of BNSL. Specifically, we construct an instance of QUBO whose solution is the score-maximizing DAG; there is a simple transformation between a classically defined QUBO instance and a diagonal quantum 2-local Hamiltonian consisting of only Pauli $Z$ and $ZZ$ terms [220].

When the desired Hamiltonian is diagonal and 2-local an embedding technique called graph-minor embedding can be used [73, 75]. A graph $G$ is a minor of another graph $H$ if there exists a mapping from vertices of $G$ to disjoint, individually connected subgraphs of $H$ such that for every edge $e$ in $G$ there is an edge in $H$ whose adjacent vertices are mapped to by the adjacent vertices of the edge $e$. The desired Hamiltonian and hardware are considered as graphs, called the logical and physical respectively, where qubits correspond to vertices and edges correspond to a 2-body interaction, desired or available. Graph-minor embedding consists of two parts: finding a mapping of the logical vertices to sets of physical as described, and setting the parameters of the physical Hamiltonian such that the logical fields are distributed among the appropriate physical qubits and there is a strong

ferromagnetic coupling between physical qubits mapped to my the same logical qubit so that they act as one. Determining the graph-minor mapping, or even if the logical graph is a minor of the physical one, is itself NP-hard, and so in practice heuristics are used [61].

## 4.3 Mapping BNSL to QUBO

We use $n(n-1)$ bits $\mathbf{d} = (d_{ij})_{\substack{1 \leq i < j \leq n \\ i \neq j}}$ to encode each of the possible arcs in a directed graph, where $d_{ij} = 1$ indicates the presence of the arc from vertex $X_i$ to vertex $X_j$ and $d_{ij} = 0$ indicates its absence. In this way, the matrix whose entries are $\{d_{ij}\}$ is the adjacency matrix of a directed graph (where $d_{ii} = 0$). Let $G(\mathbf{d})$ be that directed graph encoded in some $\mathbf{d}$. The mapping consists of the construction of a function of these "arc bits" that is equal to the logarithm of the score of the structure they encode, as well as a function that penalizes states that encode graphs with directed cycles. Additionally, due to resource constraints, we add a function that penalizes structures in which any node has more than $m$ parents and allow that the scoring function only works on states that encode structures in which each vertex has at most $m$ parents.

### 4.3.1 Score Hamiltonian

For numerical efficiency, it is the logarithm of the likelihood for a given structure that is actually computed in practice. The likelihood given in Equation 4.3 decomposes into a product of likelihoods for each variable, which we exploit here. Let

$$s_i(\Pi_i(B_S)) \equiv -\log\left(\prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(N_{ij}+\alpha_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk}+N_{ijk})}{\Gamma(\alpha_{ijk})}\right), \tag{4.5}$$

i.e. the negation of the "local" score function, and

$$s(B_S) \equiv \sum_{i=1}^{n} s_i(\Pi_i(B_S)), \tag{4.6}$$

so that

$$\log p(D|B_S) = -s(B_S) = -\sum_{i=1}^{n} s_i(\Pi_i(B_S)). \tag{4.7}$$

The negation is included because while we wish to maximize the likelihood, in QUBO the objective function is minimized. We wish to define a quadratic pseudo-Boolean function $H_{\text{score}}(\mathbf{d})$ such that $H_{\text{score}}(\mathbf{d}) = s(G(\mathbf{d}))$. Let $\mathbf{d}_i \equiv (d_{ji})_{\substack{1\leq j\leq n \\ j\neq i}}$ and define

$$H_{\text{score}}(\mathbf{d}) \equiv \sum_{i=1}^{n} H_{\text{score}}(\mathbf{d}_i). \tag{4.8}$$

Any pseudo-Boolean such as $H^{(i)}_{\text{score}}$ has a unique multinomial form and $s_i(\Pi_i(G(\mathbf{d})))$ depends only on arcs whose head is $X_i$ (i.e. those encoded in $\mathbf{d}_i$), so we write without loss of generality

$$H^{(i)}_{\text{score}}(\mathbf{d}_i) = \sum_{J\subset\{1,\cdots,n\}\setminus\{i\}} \left( w_i(J) \prod_{j\in J} d_{ji} \right). \tag{4.9}$$

From this it is clear that $w_i(\emptyset) = s_i(\emptyset)$. If $X_i$ has a single parent $X_j$, then the above simplifies to

$$H^{(i)}_{\text{score}} = w_i(\emptyset) + w_i(\{j\}) = s_i(\{X_j\}), \tag{4.10}$$

which yields $w_i(\{j\}) = s_i(\{X_j\}) - s_i(\emptyset)$ for arbitrary $j$. Similarly, if $X_i$ has two parents $X_j$ and $X_k$, then

$$\begin{aligned} H^{(i)}_{\text{score}} &= w_i(\emptyset) + w_i(\{j\}) + w_i(\{k\}) + w_i(\{j,k\}) \\ &= s_i(\emptyset) + (s_i(\{X_j\}) - s_i(\emptyset)) + (s_i(\{X_j\}) - s_i(\emptyset)) + w_i(\{j,k\}) \\ &= s_i(\{X_j\}) + s_i(\{X_k\}) - s_i(\emptyset) + w_i(\{j,k\}) \\ &= s_i(\{X_j, X_k\}), \end{aligned} \tag{4.11}$$

which yields $w_i(\{j,k\}) = s_i(\{X_j, X_k\}) - s_i(\{X_j\}) - s_i(\{X_k\}) + s_i(\emptyset)$. Extrapolating this pattern, we find that

$$w_i(J) = \sum_{l=0}^{|J|} (-1)^{|J|-l} \sum_{\substack{K\subset J \\ |K|=l}} s_i(K). \tag{4.12}$$

Note that the general form given in Equation 4.9 includes terms of order $(n-1)$. Ultimately, we require a quadratic function and reducing high-order terms to quadratic requires many extra variables. Therefore, we limit the number of parents that each variable has to $m$ via $H_{\max}$, described below, and allow that the score Hamiltonian actually gives the score only for structures with maximum in-degree $m$:

$$H^{(i)}_{\text{score}}(\mathbf{d}_i) = \sum_{\substack{J\subset\{1,\cdots,n\}\setminus\{i\} \\ |J|\leq m}} \left( w_i(J) \prod_{j\in J} d_{ji} \right), \tag{4.13}$$

which is equal to $s_i(\Pi_i(G(\mathbf{d})))$ if $|\mathbf{d}_i| \leq m$.

### 4.3.2 Max Hamiltonian

Now we define a function $H^{(i)}_{\max}$ whose value is zero if variable $X_i$ has at most $m$ parents and positive otherwise. This is done via a slack variable $y_i$ for each node. Define

$$d_i \equiv |\mathbf{d}_i| = \sum_{\substack{1\leq j\leq n \\ j\neq i}} d_{ji}, \tag{4.14}$$

i.e. $d_i$ is the in-degree of $x_i$,

$$\mu \equiv \lceil \log_2(m+1) \rceil, \tag{4.15}$$

i.e. $\mu$ is the number of bits needed to represent an integer in $[0, m]$,

$$y_i \equiv \sum_{l=1}^{\mu} 2^{l-1} y_{il}, \tag{4.16}$$

i.e. $y_i \in \mathbb{Z}$ is encoded using the $\mu$ bits $\mathbf{y}_i = (y_{il})_{l=1}^{\mu} \in \mathbb{B}^{\mu}$, and

$$H^{(i)}_{\max}(\mathbf{d}_i, \mathbf{y}_i) = \delta^{(i)}_{\max}(m - d_i - y_i)^2, \tag{4.17}$$

where $\delta^{(i)}_{\max} > 0$ is the weight of the penalty. For convenience, we also write $H^{(i)}_{\max}(d_i, y_i)$ without loss of generality. When viewed as a quadratic polynomial of $y_i$, $H^{(i)}_{\max}$ takes its

minimal value of zero when $y_i = m - d_i$. Note that $0 \leq y_i \leq 2^\mu - 1$. If $d_i \leq m$, let $y_i^*$ be such that $0 \leq y_i^* = m - d_i \leq m \leq 2^\mu - 1$. Then $H_{\max}(d_i, y_i^*) = 0$. However, when $d_i > m$, because $y_i \geq 0$, we cannot set $y_i$ in that way. By taking the derivative with respect to $y_i$,

$$\frac{\partial}{\partial y_i} H_{\max}^{(i)}(d_i, y_i) = 2\delta_{\max}^{(i)}(y_i - m + d_i) > 0, \tag{4.18}$$

we see that $H_{\max}^{(i)}$ takes its minimum value over the domain of $y_i$ when $y_i = 0$, and that value is

$$H_{\max}^{(i)}(d_i, 0) = \delta_{\max}^{(i)}(m - d_i)^2. \tag{4.19}$$

Noting that $H_{\max}^{(i)}$ is nonnegative,

$$\min_{y_i} H_{\max}^{(i)}(d_i, y_i) = \begin{cases} 0, & d_i \leq m, \\ \delta_{\max}(d_i - m)^2, & d_i > m. \end{cases} \tag{4.20}$$

Thus, if the constraint $|\mathbf{d}_i| \leq m$ is satisfied, $H_{\max}^{(i)}$ does nothing, but if $|\mathbf{d}_i| > m$, a penalty of at least $\delta_{\max}^{(i)}$ is added.

### 4.3.3 Acyclicity

Lastly, we must ensure that the structure encoded in $\{d_{ij}\}$ has no directed cycles. We do so by introducing additional Boolean variables $\mathbf{r} = (r_{ij})_{1 \leq i < j \leq n}$ that will encode a binary relation on the set of variables. Every directed acyclic graph admits at least one topological order of the vertices, and a graph with a directed cycle admits none. A topological order "$\leq$" of the vertices $\{X_i\}$ of a digraph is a total order thereon such that for every edge $(i, j)$ in the digraph $X_i \leq X_j$. Such an order is not unique in general. Let $r_{ij} = 1$ represent $x_i \leq x_j$ and $r_{ij} = 0$ represent $x_i \geq x_j$.

To ensure acyclicity, we define a function $H_{\text{trans}}(\mathbf{r})$ such that $H_{\text{trans}}(\mathbf{r})$ is zero if the relation encoded in $\{r_{ij}\}$ is transitive and is positive otherwise, as well as a function $H_{\text{consist}}$

such that $H_{\text{consist}}(\mathbf{d}$ is zero if the order encoded in $\{r_{ij}\}$ is consistent with the directed edge structure encoded by $\{d_{ij}\}$ and positive otherwise. First, we ensure that $\{r_{ij}\}$ is transitive. Because if a tournament has any cycle, it has a cycle of length three, it is sufficient to penalize directed 3-cycles. Define

$$H_{\text{trans}}(\mathbf{r}) \equiv \sum_{1 \le i < j \le n} H_{\text{trans}}^{(ijk)}(r_{ij}, r_{ik}, r_{jk}), \tag{4.21}$$

where

$$\begin{aligned} H_{\text{trans}}^{(ijk)}(r_{ij}, r_{ik}, r_{jk}) &\equiv \delta_{\text{trans}}^{(ijk)} \left[ r_{ij} r_{jk} (1 - r_{ik}) + (1 - r_{ij})(1 - r_{jk}) r_{ik} \right] \\ &= \delta_{\text{trans}}^{(ijk)} \left( r_{ik} + r_{ij} r_{jk} - r_{ij} r_{ik} - r_{jk} r_{ik} \right) \\ &= \begin{cases} \delta_{\text{trans}}^{(ijk)}, & \left[ (x_i \le x_j \le x_k \le x_i) \vee (x_i \ge x_j \ge x_k \ge x_i) \right], \\ 0, & \text{otherwise}, \end{cases} \end{aligned} \tag{4.22}$$

and $\delta_{\text{trans}}^{(ijk)}$ is the penalty weight added if $\mathbf{r}$ encodes either 3-cycle containing $\{x_i, x_j, x_k\}$. Note that the superscripted indices on the penalty weight variable are unordered so that $\delta_{\text{trans}}^{(i'j'k')} \equiv \delta_{\text{trans}}^{(ijk)}$ for all permutations $(i', j', k')$ of $(i, j, k)$.

Second, we must penalize any state that represents an order and a directed graph that are inconsistent with each other, i.e. in which $r_{ij} = 1$ and $(x_j, x_i) \in E(G(\mathbf{d}))$ or $r_{ij} = 0$ and $((x_i, x_j) \in E(G(\mathbf{d}))$. Equivalently, we want to ensure that neither $r_{ij} = d_{ji} = 1$ nor $r_{ij} = 1 - d_{ij} = 0$. Define

$$H_{\text{consist}}(\mathbf{d}, \mathbf{r}) \equiv \sum_{1 \le i < j \le n} H_{\text{consist}}^{(ij)}(d_{ij}, d_{ji}, r_{ij}) \tag{4.23}$$

and

$$H_{\text{consist}}(d_{ij}, d_{ji}, r_{ij}) = \delta^{(ij)}_{\text{consist}}(d_{ji}r_{ij} + d_{ij}(1 - r_{ij}))$$
$$= \begin{cases} \delta^{(ij)}_{\text{consist}}, & d_{ji} = r_{ij} = 1 \vee (d_{ij} = 1 \wedge r_{ij} = 0), \\ 0, & \text{otherwise}, \end{cases} \tag{4.24}$$

which has the desired features. Again the superscripted indices on the penalty weight variable are unordered, so that $\delta^{(ji)}_{\text{consist}} \equiv \delta^{(ij)}_{\text{consist}}$ for $1 \leq i < j \leq n$. Finally, define

$$H_{\text{cycle}}(\mathbf{d}, \mathbf{r}) \equiv H_{\text{consist}}(\mathbf{d}, \mathbf{r}) + H_{\text{trans}}(\mathbf{r}), \tag{4.25}$$

which takes on its minimal value of zero if $G(\mathbf{d})$ is a DAG and is strictly positive otherwise.

### 4.3.4 Total Hamiltonian

Putting together the parts of the Hamiltonian defined above, define

$$H(\mathbf{d}, \mathbf{y}, \mathbf{r}) \equiv H_{\text{score}}(\mathbf{d}) + H_{\max}(\mathbf{d}, \mathbf{y}) + H_{\text{cycle}}(\mathbf{d}, \mathbf{r}). \tag{4.26}$$

In the next section, we show lower bounds on the penalty weights therein that ensure that the ground state of the total Hamiltonian $H$ encodes the highest-scoring DAG with a maximum parent set size of $m$. The sets of variables described above have the following sizes:

$$|\{d_{ij}\}| = n(n-1),$$
$$|\{r_{ij}\}| = \frac{n(n-1)}{2}, \text{and}$$
$$|\{y_{il}\}| = n\mu = n\left\lceil \log_2(m+1) \right\rceil. \tag{4.27}$$

Furthermore, while $H_{\max}$ and $H_{\text{cycle}}$ are natively 2-local, $H_{\text{score}}$ is $m$-local. For each variable $x_i$ there are $\binom{n-1}{l}$ possible parent sets of size $l$ and the same number of corresponding $l$-local

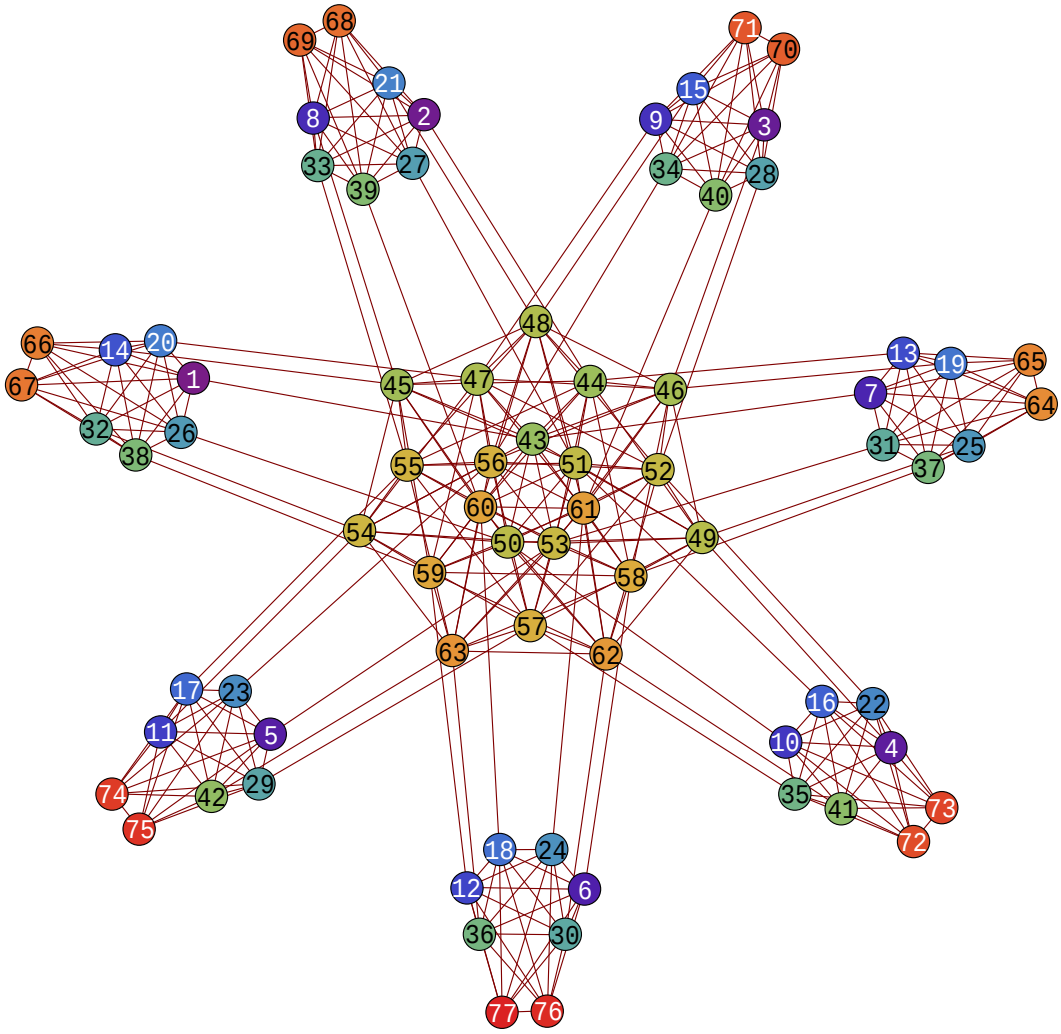



Figure 4.1: Logical Graph for $n = 7$ BN variables with a maximum of $m = 2$ parents. Each vertex corresponds to a bit in the original QUBO and an edge between two vertices indicates a non-zero quadratic term containing the corresponding bits. The central cluster is the order bits used to enforce acyclicity; it is highly connected but not complete. Each "spike corresponds to a variable $X_i$ in the Bayesian network. The outer two vertices are the corresponding slack bits $\{y_{il}\}$ and the remaining inner vertices are the arc bits $\{d_{ji}\}$ representing those arcs for which the corresponding Bayesian network variable is the head. Each spike is a clique, due to $H_{\max}$ (independent of which the arc bits for a given BN variable are fully connected due to $H_{\text{score}}$). Each arc bit is connected to a single order bit and each order bit is connected to two arc bits, due to $H_{\text{consist}}$.

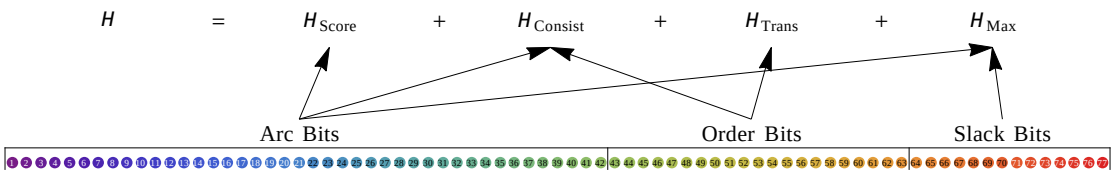



Figure 4.2: Schematic of the Hamiltonian. The row of disks represents all of the bits in the original QUBO problem, colored consistently with the logical graph above. They are grouped into three sets: the arc bits representing the presence of the possible arcs, the order bits representing a total ordering by which we enforce acyclicity, and the slack bits used to limit the size of the parent sets. An arrow from a group of bits to a part of the Hamiltonian indicates that that part of the Hamiltonian is a function of that set of bits.

terms in $H_{\text{score}}$. If $m = 3$, the full set of $\binom{n-1}{3}$ high-local terms $\left\{\prod_{j\in J} d_{ji} || J| = 3\right\}$ corresponding to parent sets of the variable $x_i$ can be reduced using $\lfloor \frac{(n-2)^2}{4} \rfloor$ ancilla variables. In total, $n\lfloor \frac{(n-2)^2}{4} \rfloor$ ancilla variables are needed to reduce $H_{\text{score}}$ to 2-local.

A quadratic pseudo-Boolean function can be identified with a graph whose vertices correspond to its arguments and whose edges correspond to non-zero quadratic terms. The graph identified with the Hamiltonian described above for $m = 2$ has several features that indicate it may be difficult to embed in sparsely connected physical devices. First, for each variable $X_i$ there is a clique consisting of the variables $\{d_{ji}\}\cup\{y_{il}\}$, whose order is $(n-1)+\mu$. Second, the set of variables $\{r_{ij}\}$ are almost fully connected.

### 4.3.5 Utilizing Prior Information

The mapping so far described assumes a uniform prior distribution over all possible DAGs of the appropriate size and with the given maximum number of parents. However, there are situations in which it may be desirable to fix the presence or absence of an arc in the search space. This could be because of domain knowledge or because hardware limitations prevent the implementation of the mapping for all arcs, in which case resort can be made to iterative search procedures such as the bootstrap method [110]. To realize the

reduction in qubits needed by accounting for such a reduced search space, suppose that we wish to only consider network structures that include the arc $(i, j)$, where $i < j$ without loss of generality. We then set $d_{ij} = 1$, $d_{ji} = 0$, and $r_{ij} = 1$. Similarly, if $(i, j)$ is to be excluded, we set $d_{ij} = d_{ji} = 0$ and keep $r_{ij}$ as a free variable. This can be done for any number of arcs. The Hamiltonian remains unchanged once these substitutions are made, and the lower bounds on the penalty weight remain sufficient, with the exception of the terms used in quadratization in the case $m > 2$, in which case the quadratization should be done after substitution to utilize the reduction in degree of some terms.

## 4.4 Penalty Weights

In the expression above, there are several sets of free parameters called penalty weights: $\{\delta^i_{\max} | 1 \leq i \leq n\}$, $\{\delta^{ij}_{\text{consist}} | 1 \leq i, j \leq n, i \neq j\}$, and $\{\delta^{ijk}_{\text{trans}} | 1 \leq i < j < k\}$. They are associated with penalty terms, i.e. parts of the Hamiltonian whose value is zero on states satisfying the corresponding constraint and is positive on states violating it. The purpose of their inclusion is to ensure that the energy-minimizing state of the total Hamiltonian satisfies the requisite constraints by increasing the energy of those that do not. More strongly, the penalty weights must be set such that the ground state of the total Hamiltonian is the lowest energy state of $H_{\text{score}}$ that satisfies the constraints. Here we provide sufficient lower bounds on the penalty weight necessary to ensure that this purpose is met. No claim is made to their necessity, and tighter lower bounds may exist. It is important to note that these bounds are mathematical, i.e. they ensure their purpose is met as stated above. In pure adiabatic quantum computation, in which the quantum system is in its ground state for the duration of the algorithm, this is sufficient (though the computation time necessary for the conditions of the adiabatic theorem to hold may be longer than otherwise if a penalized state has lower

energy than the first excited unpenalized state). In practical quantum annealing, however, a combination of physical effects may cause the optimal value (in the sense of minimizing the energy of the lowest-energy state found, which may or may not be the global ground state) of the penalty weights to be less than these bounds. This remains the case even for bounds shown to be tight with respect to their mathematical properties.

The bound presented for each of the three sets of penalty weights is based on the notion that only the addition of an arc (i.e. changing some $d_{ij}$ from 0 to 1) can lead to the violation of two of the constrains we are concerned with: the maximum number of parents and the consistency of the arc bits and the order bits. Therefore, we can use a basis for how strongly the associated penalty needs to be the greatest difference in the energy of $H_{\text{score}}$ adding each arc can contribute. The penalty for the third constraint, the absence of directed 3-cycles among the order bits, will then be a function of the penalty for the consistency constraint.

Formally, we wish to set the penalties such that for any $\mathbf{d}$ violating at least one of the constraints, we have

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d},\mathbf{y},\mathbf{r}) > H_{\text{score}}(\mathbf{d}^*), \tag{4.28}$$

where

$$\mathbf{d}^* \equiv \underset{\substack{|\mathbf{d}'|\leq m \\ G(\mathbf{d}') \text{ is a DAG}}}{\arg\min} H_{\text{score}}(\mathbf{d}'). \tag{4.29}$$

This is achieved by showing that for any such $\mathbf{d}$ violating at least one constraint, there is another $\mathbf{d}'$ that satisfies all the constraints such that

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d},\mathbf{y},\mathbf{r}) \geq \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}',\mathbf{y},\mathbf{r}). \tag{4.30}$$

Because $\mathbf{d}'$ satisfies all the constraints,

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}',\mathbf{y},\mathbf{r}) = H_{\text{score}}(\mathbf{d}'), \tag{4.31}$$

which implies the inequality in (4.28). In this section, we state the bounds and provide brief justification, but relegate the proofs to Appendix 4.6.2.

### 4.4.1 Auxiliary Quantity

In this section, we briefly define an auxiliary quantity,

$$\Delta'_{ji} = - \min_{\{d_{ki} | k \neq i,j\}} \left\{ H^{(i)}_{\text{score}} \Big|_{d_{ji}=1} - H^{(i)}_{\text{score}} \Big|_{d_{ji}=0} \right\}, \tag{4.32}$$

that will allow us to define the maximum penalty weights associated with the bounds described previously. For details of the calculation of this quantity, see Appendix 4.6.1. In general it is possible that $\Delta_{ji} < 0$ as defined above. We thus define the quantity

$$\Delta_{ji} \equiv \max\{0, \Delta'_{ji}\}. \tag{4.33}$$

In the proof of the bounds, the two following facts will be useful.

**Claim 1** (Monotonicity of $H_{\max}$)**.** *If* $\mathbf{d} \geq \mathbf{d}'$*, then* $\min_{\mathbf{y}} H_{\max}(\mathbf{d}, \mathbf{y}) \geq \min_{\mathbf{y}} H_{\max}(\mathbf{d}', \mathbf{y})$.

**Claim 2** (Monotonicity of $H_{\text{cycle}}$)**.** *If* $\mathbf{d} \geq \mathbf{d}'$*, then* $\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}, \mathbf{r}) \geq \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}', \mathbf{r})$.

These say simply that the removal of one or more arcs from $G(\mathbf{d})$ cannot increase the values of $H_{\max}$ nor $H_{\text{cycle}}$.

### 4.4.2 "Maximum" Penalty Weights

Here we show a lower bound for $\{\delta^{(i)}_{\max}\}$ that guarantees that if $\mathbf{d}$ is such that $\max_{1 \leq i \leq n} |\mathbf{d}_i| > m$ there exists a $\mathbf{d}'$ with lesser total energy such that $\max_{1 \leq i \leq n} |\mathbf{d}_i| \leq m$. To do so, we show that if, for some $\mathbf{d}$ and $i$, $|\mathbf{d}_i| > m$, there is a $\mathbf{d}'$ such that $|\mathbf{d}'_i| = |\mathbf{d}_i| - 1$ and

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) \geq \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r}). \tag{4.34}$$

This idea can be applied iteratively to show that if, for some $\mathbf{d}$ and $i$, $|\mathbf{d}_i| > m$, there is some $\mathbf{d}'$ with lesser energy such that $\mathbf{d}'_j = \mathbf{d}_j$ for $j \neq i$ and $|\mathbf{d}'_i| \leq m$. This idea in turn can be applied iteratively to show that if for some $\mathbf{d}$ $\max_{1\leq i\leq n} |\mathbf{d}_i| > m$ there is a $\mathbf{d}'$ such that $\max_{1\leq i\leq n} |\mathbf{d}'_i| \leq m$.

**Claim 3.** *If $\delta^{(i)}_{max} > \max_{j\neq i} \Delta_{ji}$ for all $1 \leq i \leq n$, then for all $\mathbf{d}$ such that, for some $i^*$, $|\mathbf{d}_{i^*}| > m$, there is a $\mathbf{d}'$ such that $|\mathbf{d}'_{i^*}| = |\mathbf{d}_{i^*}| - 1$, $\mathbf{d}'_i = \mathbf{d}_i$ for all $i \neq i^*$, and $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r})$.*

**Claim 4** (Sufficiency of "Maximum" Penalty Weight)**.** *If $\delta^{(i)}_{max} > \max_{j\neq i} \Delta_{ji}$ for all $i$, then for all $\mathbf{d}$ such that $\max_i |\mathbf{d}_i| > m$, there is a $\mathbf{d}' \leq \mathbf{d}$ such that $\max_i |\mathbf{d}'_i| \leq m$ and $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r})$.*

### 4.4.3 "Reduction" Penalty Weights

The degree of the "score" Hamiltonian $H_{\text{score}}$ is natively $m$-local as constructed. If $m = 2$, as it often will be in practice, the total Hamiltonian is natively quadratic. If $m > 2$, additional ancilla bits are needed to reduce the locality. The general method for doing this is to replace the conjunction of a pair bits with an ancilla bit and to add a penalty term with sufficiently strong weighting that penalizes states in which the ancillary bit is not equal to the conjunction to which it should be. For $m = 3$, this can be done using $n \left\lfloor (n-2)^2/4 \right\rfloor$ ancilla bits, but no more, where each $H^{(i)}_{\text{score}}$ containing $n-1$ arc bits is quadratized independently; furthermore, heuristic methods have been developed that reduce needed weight of the penalty terms [16]. For $m = 4$, at most $n\binom{n-1}{2}$ ancilla bits are needed. More generally, $\mathcal{O}(n^{2\log d})$ is ancilla bits are needed [49]. Because the proof of the bounds on the other penalty weights are secular as to the degree of $H_{\text{score}}$ so long as $\{\Delta_{ij}\}$ is computed appropriately, the quadratization of $H_{\text{score}}$, including the addition of penalty

terms and the needed weights therefor, can be done using the standard methods described in the literature independent of the other penalties described here.

### 4.4.4 "Cycle" Penalty Weights

First, we that if the consistency penalty is set high enough, for any $\mathbf{d}$ encoding a graph with a 2-cycle, there is some $\mathbf{d}'$ encoding one whose minimal value of $H$ over $\mathbf{y}, \mathbf{r}$ is strictly less than that of $\mathbf{d}$.

**Claim 5** (Removal of 2-cycles.)**.** *If* $\delta^{(ij)}_{consist} > \max\{\Delta_{ij}, \Delta_{ji}\}$ *for all* $1 \leq i < j \leq n$ *, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *contains a 2-cycle, there is some* $\mathbf{d}' \leq \mathbf{d}$ *such that* $G(\mathbf{d}')$ *does not contain a 2-cycle and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r})$.

Second, we show that for any $\mathbf{d}$ that encodes a digraph without a 2-cycle, the minimal value of $H_{\text{consist}}$ over all $\mathbf{r}$ is zero.

**Claim 6** (Sufficiency of "Consistency" Penalty Weights)**.** *If* $\delta^{(ij)}_{\text{consist}} > (n-2) \max_{k \notin \{i,j\}} \delta^{(ijk)}_{\text{trans}}$ *for* $1 \leq i < j \leq n$*, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *contains no 2-cycle,* $H_{\text{consist}}(\mathbf{d}, \mathbf{r}^*) = 0$*, where* $\mathbf{r}^* = \arg\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}, \mathbf{r})$.

Third, we show that for any $\mathbf{d}$ that encodes a digraph not containing a 2-cycle but that is not a DAG, there is some $\mathbf{d}'$ that does encode a DAG and whose minimal value of $H$ over all $\mathbf{y}, \mathbf{r}$ is strictly less than that of $\mathbf{d}$.

**Claim 7** (Sufficiency of "Transitivity" Penalty Weights)**.** *If* $\delta^{(ij)}_{\text{consist}} > (n-2) \max_{k \notin \{i,j\}} \delta^{(ijk)}_{\text{trans}}$ *for* $1 \leq i < j \leq n$ *and* $\delta^{(ijk)}_{\text{trans}} = \delta_{\text{trans}} > \max_{\substack{1 \leq i', j' \leq n \\ i' \neq j'}} \Delta_{i'j'}$ *for* $1 \leq i < j < k \leq n$*, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *does not contain a 2-cycle but does contain a directed cycle there is some* $\mathbf{d}'$ *such that* $G(\mathbf{d}')$ *is a DAG and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r})$.

Lastly, we show that for all $\mathbf{d}$ that encode a digraph that is not a DAG, there is some $\mathbf{d}'$ that does encode a DAG and whose minimal value of $H$ over all $\mathbf{y}, \mathbf{r}$ is strictly less than that of $\mathbf{d}$.

**Claim 8** (Sufficiency of "Cycle" Penalty Weights)**.** *If* $\delta_{\text{consist}}^{(ij)} > (n-2) \max_{k \notin \{i,j\}} \delta_{\text{trans}}^{(ijk)}$ *for all* $1 \leq i < j \leq n$ *and* $\delta_{\text{trans}}^{(ijk)} = \delta_{\text{trans}} > \max_{\substack{1 \leq i,j \leq n \\ i \neq j}}$ *for all* $1 \leq i < j < k \leq n$*, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *contains a directed cycle, there is a* $\mathbf{d}' \leq \mathbf{d}$ *such that* $G(\mathbf{d}')$ *is a DAG, and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r}) < \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r})$.

### 4.4.5 Overall Sufficiency

Finally, we show that the digraph encoded in the ground state of the total Hamiltonian $H$ is a DAG and has a maximum parent set size of at most $m$, and that it is the solution to the corresponding BNSL instance.

**Claim 9** (Overall Sufficiency)**.** *If* $\delta_{\max}^{(i)} > \max_{j \neq i} \Delta_{ji}$ *for all* $1 \leq i \leq n$*,* $\delta_{\text{consist}}^{(ij)} > (n-2) \max_{k \notin \{i,j\}} \delta_{\text{trans}}^{(ijk)}$ *for all* $1 \leq i < j \leq n$ *and* $\delta_{\text{trans}}^{(ijk)} = \delta_{\text{trans}} > \max_{\substack{1 \leq i',j' \leq n \\ i' \neq j'}} \Delta_{i'j'}$ *for all* $1 \leq i < j < k \leq n$*, then* $H(\mathbf{d}^*, \mathbf{y}, \mathbf{r}) = \min_{\substack{\max_i |\mathbf{d}_i| \leq m \\ G(\mathbf{d}) \text{ is a DAG}}} H_{\text{score}}(\mathbf{d}, \mathbf{y}, \mathbf{r})$*,* $G(\mathbf{d}^*)$ *is a DAG, and* $\max_i |\mathbf{d}_i^*| \leq m$*, where* $\mathbf{d}^* = \arg\min_{\mathbf{d}} \{\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r})\}$.

The strict inequalities used in the specification of the lower bounds ensures that the global ground state is a score-maximizing DAG with maximum parent set size $m$, but replacing them with weak inequalities is sufficient to ensure that the ground state energy is the greatest score over all DAGs with maximum parent set size $m$. However, the latter is of little interest in the present situation because it is the DAG itself that is of interest, not its score per se.

## 4.5 Conclusion

We have introduced a mapping from the native formulation of BNSL to QUBO that enables the solution of the former using novel methods.

The mapping is unique amongst known mappings of optimization problems to QUBO in that the logical structure is instance-independent for a given problem size. This enables the expenditure of considerably more computational resources on the problem of embedding the logical structure into a physical device because such an embedding need only be done once and reused for new instances. The problem addressed, BNSL, is special among optimization problems in that approximate solutions thereto often have value rivaling that of the exact solution. This property, along with the general intractability of exact solution, implies the great value of efficient heuristics such as SA or QA implemented using this mapping.

At present, only problems of up to seven BN variables can be embedded in existing quantum annealing hardware (i.e. the D-Wave Two chip installed at NASA Ames Research Center), whereas classical methods are able to deal with many of tens of BN variables. Nevertheless, the quantum state of the art is quickly advancing, and it is conceivable that quantum annealing could be competitively applied to BNSL in the near future. Given the already advanced state of classical simulated annealing code, it is similarly conceivable that its application to the QUBO form described here could be competitive with other classical methods for solving BNSL.

## Acknowledgements

This work was supported by the AFRL Information Directorate under grant F4HBKC4162G001. All opinions, findings, conclusions, and recommendations expressed in this material are those of the authors and do not necessarily reflect the views of AFRL. The authors would also like to acknowledge support from the NASA Advanced Exploration Systems program and NASA Ames Research Center. R. B. and A. A.-G. were supported by the National Science Foundation under award NSF CHE-1152291. The authors are grateful to David Tempel, Ole Mengshoel, and Eleanor Rieffel for useful discussions.

## 4.6 Appendix

### 4.6.1 Calculation of $\Delta_{ij}$

Recall the definition of the auxillary quantity,

$$\Delta'_{ji} = -\min_{\{d_{ki}|k\neq i,j\}} \left\{ H^{(i)}_{\text{score}}\Big|_{d_{ji}=1} - H^{(i)}_{\text{score}}\Big|_{d_{ji}=0} \right\}, \tag{4.32}$$

Note that $H^{(i)}_{\text{score}}$ can be decomposed as

$$\begin{aligned} H^{(i)}_{\text{score}} &= \sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i\} \\ |J|\leq m}} \left( w_i(J) \prod_{k\in J} d_{ki} \right) \\ &= \sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\} \\ |J|\leq m}} \left( w_i(J) \prod_{k\in J} d_{ki} \right) + \sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\} \\ |J|\leq m-1}} \left( w_i(J\cup\{j\}) d_{ji} \prod_{k\in J} d_{ki} \right), \end{aligned} \tag{4.35}$$

where the first term is independent of $d_{ji}$ and thus cancels in the argument of the minimization on the right-hand side of Equation 4.32. Thus $\Delta_{ji}$ simplifies to

$$\begin{aligned}\Delta'_{ji} &= -\min_{\{d_{ki}|k\neq i,j\}}\left\{\sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\}\\|J|\leq m-1}}\left(w_i(J\cup\{j\})\prod_{k\in J}d_{ki}\right)\right\}\\ &= \max_{\{d_{ki}|k\neq i,j\}}\left\{-\sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\}\\|J|\leq m-1}}\left(w_i(J\cup\{j\})\prod_{k\in J}d_{ki}\right)\right\}. \qquad (4.36)\end{aligned}$$

For $m = 1$, $\Delta_{ji}$ is trivially $-w_i(\{j\})$, the constant value of the expression to be extremized in Equation 4.36 regardless of the values of $\{d_{ki}|k \neq i, j\}$. For $m = 2$, $\Delta_{ji}$ can still be calculated exactly:

$$\begin{aligned}\Delta'_{ji} &= \max_{\{d_{ki}|k\neq i,j\}}\left\{-\sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\}\\|J|\leq 1}}\left(w_i(J\cup\{j\})\prod_{k\in J}d_{ki}\right)\right\}\\ &= \max_{\{d_{ki}|k\neq i,j\}}\left\{-w_i(\{j\})-\sum_{\substack{1\leq k\leq n\\k\neq i,j}}d_{ki}\right\}\\ &= -w_i(\{j\})-\sum_{\substack{1\leq k\leq n\\k\neq i,j\\w_i\{j,k\})<0}}w_i(\{j,k\})\\ &= -w_i(\{j\})-\sum_{\substack{1\leq k\leq n\\k\neq i,j}}\min\{0,w_i(\{j,k\})\}. \qquad (4.37)\end{aligned}$$

However, for $m \geq 3$, calculation of the extremum in Equation 4.36 is an intractable optimization problem in its own right and therefore we must resort to a reasonable bound. Because $\Delta_{ji}$ will be used in finding a lower bound on the necessary penalty weights, caution ditates that we use, if needed, a greater value than necessary. A reasonable upper bound

on the true value is:

$$\Delta'_{ji} = \max_{\{d_{ki}|k\neq i,j\}} \left\{ - \sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\} \\ |J|\leq m-1}} \left( w_i(J\cup\{j\}) \prod_{k\in J} d_{ki} \right) \right\}$$

$$\leq - \sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\} \\ |J|\leq m-1 \\ w_i(J\cup\{j\})<0}} w_i(J\cup\{j\}) = - \sum_{\substack{J\subset\{1,\ldots,n\}\setminus\{i,j\} \\ |J|\leq m-1}} \min\{0, w_i(J\cup\{j\})\}. \tag{4.38}$$

### 4.6.2 Proofs of Penalty Weight Lower Bounds

**Claim 1** (Monotonicity of $H_{\max}$)**.** *If* $\mathbf{d}\geq\mathbf{d}'$*, then* $\min_{\mathbf{y}} H_{\max}(\mathbf{d},\mathbf{y}) \geq \min_{\mathbf{y}} H_{\max}(\mathbf{d}',\mathbf{y})$.

*Proof.* $\mathbf{d}\geq\mathbf{d}'$ implies $\mathbf{d}_i \geq \mathbf{d}'_i$ and thus $|\mathbf{d}_i| \geq |\mathbf{d}'_i|$ for $1\leq i\leq n$. By design, $\min_{\mathbf{y}_i} H_{\max}(\mathbf{d}_i,\mathbf{y}_i) = \delta^{(i)}_{\max}\max\{0,|\mathbf{d}_i|-m\}$. Let $\mathbf{y}^* \equiv \arg\min_{\mathbf{y}} H(\mathbf{d},\mathbf{y})$. Then

$$\begin{aligned}
\min_{\mathbf{y}} H_{\max}(\mathbf{d},\mathbf{y}) &= \sum_{i=1}^{n} \min_{\mathbf{y}_i} H_{\max}(\mathbf{d}_i,\mathbf{y}_i) = \sum_{i=1}^{n} \delta^{(i)}_{\max}\max\{0,|\mathbf{d}_i|-m\} \\
&\geq \sum_{i=1}^{n} \delta^{(i)}_{\max}\max\{0,|\mathbf{d}'_i|-m\} \\
&= \sum_{i=1}^{n} \min_{\mathbf{y}_i} H_{\max}(\mathbf{d}'_i,\mathbf{y}_i) = \min_{\mathbf{y}} H_{\max}(\mathbf{d}',\mathbf{y}).
\end{aligned} \tag{4.39}$$

□

**Claim 2** (Monotonicity of $H_{\text{cycle}}$)**.** *If* $\mathbf{d}\geq\mathbf{d}'$*, then* $\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d},\mathbf{r}) \geq \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}',\mathbf{r})$.

*Proof.* In the statement of the claim, we implicitly assume that $\delta^{(ij)}_{\text{consist}} > 0$ for all $1\leq i < j\leq n$. Let $\mathbf{r}^* = \arg\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d},\mathbf{r})$. Because $d_{ji}\geq d'_{ji}$ for all $i,j$ such that$i\neq j$ and

$1 \leq i, j \leq n$, and because $0 \leq r_{ij} \leq 1$ for all $1 \leq i < j \leq n$,

$$\begin{aligned}
\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}, \mathbf{r}) &= \min_{\mathbf{r}} \{H_{\text{consist}}(\mathbf{d}, \mathbf{r}) + H_{\text{trans}}(\mathbf{r})\} = H_{\text{consist}}(\mathbf{d}, \mathbf{r}^*) + H_{\text{trans}}(\mathbf{r}^*) \\
&= \left( \sum_{1 \leq i < j \leq n} \delta_{\text{consist}}^{(ij)} \left[ d_{ji} r_{ij}^* + d_{ij}(1 - r_{ij}^*) \right] \right) + H_{\text{trans}}(\mathbf{r}^*) \\
&\geq \left( \sum_{1 \leq i < j \leq n} \delta_{\text{consist}}^{(ij)} \left[ d'_{ji} r_{ij}^* + d'_{ij}(1 - r_{ij}^*) \right] \right) + H_{\text{trans}}(\mathbf{r}^*) \\
&= H_{\text{consist}}(\mathbf{d}', \mathbf{r}^*) + H_{\text{trans}}(\mathbf{r}^*) \\
&\geq \min_{\mathbf{r}} \left[ H_{\text{consist}}(\mathbf{d}', \mathbf{r}) + H_{\text{trans}}(\mathbf{r}) \right] = \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}', \mathbf{r}) \qquad (4.40)
\end{aligned}$$

□

**Claim 3.** *If $\delta_{max}^{(i)} > \max_{j \neq i} \Delta_{ji}$ for all $1 \leq i \leq n$, then for all $\mathbf{d}$ such that, for some $i^*$, $|\mathbf{d}_{i^*}| > m$, there is a $\mathbf{d}'$ such that $|\mathbf{d}'_{i^*}| = |\mathbf{d}_{i^*}| - 1$, $\mathbf{d}'_i = \mathbf{d}_i$ for all $i \neq i^*$, and $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r})$.*

*Proof.* We prove the existence of such a $\mathbf{d}'$ by construction. Let $\mathbf{d}'_i \equiv \mathbf{d}_i$ for all $i \neq i^*$. Let $\mathbf{d}'_{i^*} \equiv \mathbf{d}_{i^*}|_{d_{j^* i^*} = 0}$, where $j^* = \arg\min_{j \in \{j | d_{ji^*} = 1\}} \Delta_{ji^*}$. First, we note that by design $\min_{\mathbf{y}_i} H_{\max}^{(i)}(\mathbf{d}_i, \mathbf{y}_i) = \min\{0, \delta_{\max}^{(i)}(|\mathbf{d}_i| - m)\}$. Thus

$$\begin{aligned}
\min_{\mathbf{y}_{i^*}} H_{\max}^{(i^*)}(\mathbf{d}_{i^*}, \mathbf{y}_{i^*}) - \min_{\mathbf{y}_{i^*}} H_{\max}^{(i^*)}(\mathbf{d}'_{i^*}, \mathbf{y}_{i^*}) &= \left[ \delta_{\max}^{(i^*)}(|\mathbf{d}| - m) \right] + \left[ \delta_{\max}^{(i^*)}(|\mathbf{d}'| - m) \right] \\
&= \delta_{\max}^{(i^*)}(|\mathbf{d}| - |\mathbf{d}'|) = \delta_{\max}^{(i^*)} \\
&> \max_{j \neq i^*} \Delta_{ji^*} \geq H_{\text{score}}^{(i^*)}(\mathbf{d}'_{i^*}) - H_{\text{score}}^{(i^*)}(\mathbf{d}_{i^*}). \qquad (4.41)
\end{aligned}$$

which rearranges to

$$H_{\text{score}}^{(i^*)}(\mathbf{d}_{i^*}) + \min_{\mathbf{y}_{i^*}} H_{\max}^{(i^*)}(\mathbf{d}_{i^*}, \mathbf{y}_{i^*}) > H_{\text{score}}^{(i^*)}(\mathbf{d}'_{i^*}) + \min_{\mathbf{y}_{i^*}} H_{\max}^{(i^*)}(\mathbf{d}'_{i^*}, \mathbf{y}_{i^*}). \qquad (4.42)$$

In context,

$$\begin{aligned}
&\min_{\mathbf{y}} \left[H_{\text{score}}(\mathbf{d}) + H_{\max}(\mathbf{d}, \mathbf{y})\right] = H_{\text{score}}(\mathbf{d}) + \min_{\mathbf{y}} H_{\max}(\mathbf{d}, \mathbf{y}) \\
&= \sum_{i \neq i^*} \left[ H^{(i)}_{\text{score}}(\mathbf{d}_i) + \min_{\mathbf{y}_i} H^{(i)}_{\max}(\mathbf{d}_i, \mathbf{y}_i) + H^{(i^*)}_{\text{score}}(\mathbf{d}_{i^*}) + \min_{\mathbf{y}_{i^*}} H^{(i^*)}_{\max}(\mathbf{d}_{i^*}, \mathbf{y}_{i^*}) \right] \\
&= \sum_{i \neq i^*} \left[ H^{(i)}_{\text{score}}(\mathbf{d}'_i) + \min_{\mathbf{y}_i} H^{(i)}_{\max}(\mathbf{d}'_i, \mathbf{y}_i) + H^{(i^*)}_{\text{score}}(\mathbf{d}_{i^*}) + \min_{\mathbf{y}_{i^*}} H^{(i^*)}_{\max}(\mathbf{d}_{i^*}, \mathbf{y}_{i^*}) \right] \\
&> \sum_{i \neq i^*} \left[ H^{(i)}_{\text{score}}(\mathbf{d}'_i) + \min_{\mathbf{y}_i} H^{(i)}_{\max}(\mathbf{d}'_i, \mathbf{y}_i) + H^{(i^*)}_{\text{score}}(\mathbf{d}'_{i^*}) + \min_{\mathbf{y}_{i^*}} H^{(i^*)}_{\max}(\mathbf{d}'_{i^*}, \mathbf{y}_{i^*}) \right] \\
&= H_{\text{score}}(\mathbf{d}') + \min_{\mathbf{y}} H_{\max}(\mathbf{d}', \mathbf{y}) = \min_{\mathbf{y}} \left[H_{\text{score}}(\mathbf{d}') + H_{\max}(\mathbf{d}', \mathbf{y})\right]. \qquad (4.43)
\end{aligned}$$

By Claim 1 and the fact that $\mathbf{d}' \leq \mathbf{d}$,

$$\min_{\mathbf{r}} H_{\max}(\mathbf{d}, \mathbf{r}) \geq \min_{\mathbf{r}} H_{\max}(\mathbf{d}', \mathbf{r}), \qquad (4.44)$$

and so

$$\begin{aligned}
&\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) = H_{\text{score}}(\mathbf{d}) + \min_{\mathbf{y}} H_{\max}(\mathbf{d}, \mathbf{y}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}, \mathbf{r}) \\
&> H_{\text{score}}(\mathbf{d}') + \min_{\mathbf{y}} H_{\max}(\mathbf{d}', \mathbf{y}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}', \mathbf{r}) = \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}). \qquad (4.45)
\end{aligned}$$

□

**Claim 4** (Sufficiency of "Maximum" Penalty Weight)**.** *If* $\delta^{(i)}_{max} > \max_{j \neq i} \Delta_{ji}$ *for all* $i$*, then for all* $\mathbf{d}$ *such that* $\max_i |\mathbf{d}_i| > m$*, there is a* $\mathbf{d}' \leq \mathbf{d}$ *such that* $\max_i |\mathbf{d}'_i| \leq m$ *and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r})$*.*

*Proof.* We prove the sufficiency of the given bound by iterative application of Claim 3. Let $\mathbf{d}^{(0,0)} \equiv \mathbf{d}$. For all $i$, if $|\mathbf{d}_i| > m$, let $\mathbf{d}^{(i,|\mathbf{d}_i|-m)} \equiv \mathbf{d}^{(i-1,0)}$, and if $|\mathbf{d}_i| \leq m$, let $\mathbf{d}^{(i,0)} \equiv \mathbf{d}^{(i-1,0)}$. For all $1 \leq i \leq n$ and $x$ such that $1 \leq i \leq n$ and $1 \leq x \leq \max\{0, |\mathbf{d}_i| - m\}\}$, $|\mathbf{d}^{(i,x)}_i| > m$ and so by Claim 3 there is a $\mathbf{d}^{(i,x-1)} \leq \mathbf{d}^{(i,x)}$ such that $|\mathbf{d}^{(i,x-1)}| = |\mathbf{d}^{(i,x)}| - 1$ and

$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(i,x-1)},\mathbf{y},\mathbf{r}) < \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(i,x)},\mathbf{y},\mathbf{r})$. Then for all $i$, if $|\mathbf{d}_i| > m$, there is a sequence $\mathbf{d}^{(i,|\mathbf{d}_i|-m)},\ldots,\mathbf{d}^{(i,x)},\ldots,\mathbf{d}^{(i,0)}$ such that $\mathbf{d}^{(i,0)} \leq \mathbf{d}^{(i,|\mathbf{d}_i|-m)}$ and $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(i,0)},\mathbf{y},\mathbf{r}) < \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(i,|\mathbf{d}_i|-m)},\mathbf{y},\mathbf{r})$. Similarly, for all $i$, $\mathbf{d}^{(i,0)} \leq \mathbf{d}^{(i-1,0)}$ and $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(i,0)} \leq \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(i-1)},\mathbf{y},\mathbf{r})$, with strict inequality if $|\mathbf{d}_i| > 0$ and equality if $|\mathbf{d}_i| = 0$. Thus there is a sequence $\mathbf{d}^{(0,0)},\mathbf{d}^{(1,0)},\ldots,\mathbf{d}^{(i,0)},\ldots,\mathbf{d}^{(n,0)}$ such that $\mathbf{d}^{(n,0)} \leq \mathbf{d}^{(0,0)} = \mathbf{d}$, $\max_i |\mathbf{d}_i^{(n,0)}| \leq m$, and $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}',\mathbf{y},\mathbf{r}) < \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d},\mathbf{y},\mathbf{r})$. Setting $\mathbf{d}' \equiv \mathbf{d}^{(n,0)}$ completes the proof. □

**Claim 5** (Removal of 2-cycles.)**.** *If* $\delta_{consist}^{(ij)} > \max\{\Delta_{ij},\Delta_{ji}\}$ *for all* $1 \leq i < j \leq n$ *, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *contains a 2-cycle, there is some* $\mathbf{d}' \leq \mathbf{d}$ *such that* $G(\mathbf{d}')$ *does not contain a 2-cycle and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d},\mathbf{y},\mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}',\mathbf{y},\mathbf{r})$.

*Proof.* Let $\mathbf{d}^{(0)} \equiv \mathbf{d}$ and $l^*$ be the number of 2-cycles contained in $G(\mathbf{d})$. The claim is proved iteratively by showing that for all $\mathbf{d}^{(l)}$ such that $G(\mathbf{d}^{(l)})$ contains a 2-cycle there exists some $\mathbf{d}^{(l+1)}$ such that $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(l)},\mathbf{y},\mathbf{r}) > \min_{\mathbf{y},\mathbf{r}}(\mathbf{d}^{(l+1)},\mathbf{y},\mathbf{r})$ and $G(\mathbf{d}^{(l+1)})$ contains one fewer 2-cycle. Because a graph of fixed order can only have a finite number of 2-cycles, this implies the existence of a sequence $\mathbf{d},\mathbf{d}^{(1)},\ldots,\mathbf{d}^{(l)},\ldots,\mathbf{d}^{(l^*)}$ such that $\mathbf{d}^{(l^*)}$ meets the desiderata.

Consider an arbitrary $\mathbf{d}^{(l)}$. If $G(\mathbf{d}^{(l)})$ does not contain a directed 2-cycle, then $l = l^*$ and so we set $\mathbf{d}' = \mathbf{d}^{(l^*)}$ to complete the proof. Otherwise, choose some 2-cycle in $G(\mathbf{d}^{(l)})$ arbitrarily, i.e. some pair $\{i,j\}$ such that $(i,j),(j,i) \in E(G(\mathbf{d}))$, or, equivalently, that $d_{ij} = d_{ji} = 1$. Without loss of generality, assume $i < j$. Let $\mathbf{r}^* \equiv \arg\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)},\mathbf{r})$ and

$$(i^*,j^*) \equiv \begin{cases} (j,i), & r_{ij}^* = 1, \\ (i,j), & r_{ij}^* = 0, \end{cases} \tag{4.46}$$

i.e. the arc in $G(\mathbf{d}^{(l)}$ inconsistent with $G(\mathbf{r}^*)$. Define $\mathbf{d}^{(l+1)}$ such that

$$d_{ij}^{(l+1)} = \begin{cases} d_{ij}, & (i,j) \neq (i^*, j^*) \\ 0 = 1 - d_{ij}, & (i,j) = (i^*, j^*) \end{cases}. \tag{4.47}$$

Thus $\mathbf{d}^{(l+1)} \leq \mathbf{d}^{(l)}$ and $|\mathbf{d}^{(l+1)}| = |\mathbf{d}^{(l)}| - 1$. By construction, $G(\mathbf{d}^{(l+1)})$ contains one fewer 2-cycle than $G(\mathbf{d}^{(l)})$. Furthermore,

$$\begin{aligned}
&\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}) - \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l+1)}, \mathbf{r}) = H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}^*) - \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l+1)}, \mathbf{r}) \\
&\geq H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}^*) - H_{\text{cycle}}(\mathbf{d}^{(l+1)}, \mathbf{r}^*) \\
&= \left( \sum_{1 \leq i < j \leq n} \delta_{\text{consist}}^{(ij)} \left[ d_{ji}^{(l)} r_{ij}^* + d_{ij}^{(l)} (1 - r_{ij}^*) \right] \right) - \left( \sum_{1 \leq i < j \leq n} \delta_{\text{consist}}^{(ij)} \left[ d_{ji}^{(l+1)} r_{ij}^* + d_{ij}^{(l+1)} (1 - r_{ij}^*) \right] \right) \\
&= \sum_{1 \leq i < j \leq n} \delta_{\text{consist}}^{(ij)} \left[ (d_{ji}^{(l)} - d_{ji}^{(l+1)}) r_{ij}^* + (d_{ij}^{(l)} - d_{ij}^{(l+1)}) (1 - r_{ij}^*) \right] \\
&= \begin{cases} \delta_{\text{consist}}^{(j^* i^*)} (d_{i^* j^*}^{(l)} - d_{i^* j^*}^{(l+1)}) r_{j^* i^*}^*, & j^* < i^* \\ \delta_{\text{consist}}^{(i^* j^*)} (d_{i^* j^*}^{(l)} - d_{i^* j^*}^{(l+1)}) (1 - r_{i^* j^*}^*), & i^* < j^* \end{cases} \\
&= \begin{cases} \delta_{\text{consist}}^{(j^* i^*)}, & j^* < i^* \\ \delta_{\text{consist}}^{(i^* j^*)}, & i^* < j^* \end{cases} \\
&> \Delta_{i^* j^*} \geq H_{\text{score}}(\mathbf{d}^{(l+1)}) - H_{\text{score}}(\mathbf{d}^{(l)}).
\end{aligned} \tag{4.48}$$

By Claim 1 and the fact that $\mathbf{d}^{(l+1)} \leq \mathbf{d}^{(l)}$,

$$\min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l)}) \geq \min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l+1)}). \tag{4.49}$$

Thus,

$$\begin{aligned}
\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(l)}, \mathbf{y}, \mathbf{r}) - \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(l+1)}, \mathbf{y}, \mathbf{r}) &= H_{\text{score}}(\mathbf{d}^{(l)}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}) \min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l)}, \mathbf{y}) \\
&\quad - H_{\text{score}}(\mathbf{d}^{(l+1)}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l+1)}, \mathbf{r}) + \min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l+1)}, \mathbf{y})
\end{aligned}$$

$$\geq \left(H_{\text{score}}(\mathbf{d}^{(l)}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r})\right) - \left(H_{\text{score}}(\mathbf{d}^{(l+1)}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l+1)}, \mathbf{r})\right) > 0, \tag{4.50}$$

which rearranges to the desired inequality. □

**Claim 6** (Sufficiency of "Consistency" Penalty Weights)**.** *If* $\delta^{(ij)}_{\text{consist}} > (n-2)\max_{k\notin\{i,j\}} \delta^{(ijk)}_{\text{trans}}$ *for* $1 \leq i < j \leq n$*, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *contains no 2-cycle,* $H_{\text{consist}}(\mathbf{d}, \mathbf{r}^*) = 0$*, where* $\mathbf{r}^* = \arg\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}, \mathbf{r})$*.*

*Proof.* We prove the claim via its contrapositive: for all $\mathbf{d}, \mathbf{r}$, if $H_{\text{consist}}(\mathbf{d}, \mathbf{r}) > 0$, there is some $\mathbf{r}'$ such that $H_{\text{cycle}}(\mathbf{d}, \mathbf{r}) > H_{\text{cycle}}(\mathbf{d}, \mathbf{r}')$, so $\mathbf{r} \neq \arg\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}, \mathbf{r})$.

Consider an arbitrary $\mathbf{d}$ and some $\mathbf{r}$ such that $H_{\text{consist}}(\mathbf{d}, \mathbf{r}) > 0$. The positivity of $H_{\text{consist}}(\mathbf{d}, \mathbf{r})$ indicates that there is at least one inconsistency between $\mathbf{d}$ and $\mathbf{r}$, i.e. there is some $(i^*, j^*)$ such that $d_{i^*j^*} = \begin{cases} r_{j^*i^*}, & i^* > j^* \\ 1 - r_{i^*j^*}, & i^* < j^* \end{cases} = 1$. For convenience, we prove the claim for the case in which $i^* < j^*$; the proof provided can be easily modified for the case in which $i^* > j^*$. Let $\mathbf{r}'$ be the same as $\mathbf{r}$ exept in the bit corresponding to this inconsistency: $r'_{ij} \equiv \begin{cases} r_{ij} & (i,j) \neq (i^*, j^*) \\ 1 - r_{ij}, & (i,j) = (i^*, j^*) \end{cases}$. Then

$$\begin{aligned}
H_{\text{consist}}(\mathbf{d}, \mathbf{r}) - H_{\text{consist}}(\mathbf{d}, \mathbf{r}') &= \sum_{1\leq i<j\leq n} \left[H^{(ij)}_{\text{consist}}(d_{ij}, d_{ji}, r_{ij}) - H^{(ij)}_{\text{consist}}(d_{ij}, d_{ji}, r'_{ij})\right] \\
&= H^{(i^*j^*)}_{\text{consist}}(d_{i^*j^*}, d_{j^*i^*}, r_{i^*j^*}) - H^{(i^*j^*)}_{\text{consist}}(d_{i^*j^*}, d_{j^*i^*}, r'_{i^*j^*}) \\
&= \delta^{(i^*j^*)}_{\text{consist}}\left[d_{j^*i^*} r_{i^*j^*} + d_{i^*j^*}(1 - r_{i^*j^*})\right] - \delta^{(i^*j^*)}_{\text{consist}}\left[d_{j^*i^*} r'_{i^*j^*} + d_{i^*j^*}(1 - r'_{i^*j^*})\right] \\
&= \delta^{(i^*j^*)}_{\text{consist}}\left[-d_{j^*i^*} + d_{i^*j^*}\right] = \delta^{(i^*j^*)}_{\text{consist}}.
\end{aligned} \tag{4.51}$$

Furthermore,

$$H_{\text{trans}}(\mathbf{r}) - H_{\text{trans}}(\mathbf{r}') = \sum_{1\leq i<j<k\leq n} H^{(ijk)}_{\text{trans}}(r_{ij}, r_{ik}, r_{jk}) - \sum_{1\leq i<j<k\leq n} H^{(ijk)}_{\text{trans}}(r'_{ij}, r'_{ik}, r'_{jk})$$

$$= \sum_{\substack{1\le i<j<k\le n \\ \{i^*,j^*\}\subset\{i,j,k\}}} \Big[H^{(ijk)}_{\text{trans}}(r_{ij},r_{ik},r_{jk}) - H^{(ijk)}_{\text{trans}}(r'_{ij},r'_{ik},r'_{jk})\Big]$$

$$= \sum_{k<i^*} \delta^{(ki^*j^*)}_{\text{trans}} \big[\, (r_{kj^*} + r_{ki^*}r_{i^*j^*} - r_{ki^*}r_{kj^*} - r_{i^*j^*}r_{kj^*}) - \big(r'_{kj^*} + r'_{ki^*}r'_{i^*j^*} - r'_{ki^*}r'_{kj^*} - r'_{i^*j^*}r'_{kj^*}\big) \,\big]$$

$$+ \sum_{i^*<k<j^*} \delta^{(i^*kj^*)}_{\text{trans}} \big[\, (r_{i^*j^*} + r_{i^*k}r_{kj^*} - r_{i^*k}r_{i^*j^*} - r_{kj^*}r_{i^*j^*}) - \big(r'_{i^*j^*} + r'_{i^*k}r'_{kj^*} - r'_{i^*k}r'_{i^*j^*} - r'_{kj^*}r'_{i^*j^*}\big) \,\big]$$

$$+ \sum_{j^*<k} \delta^{(i^*j^*k)}_{\text{trans}} \big[\, (r_{i^*k} + r_{i^*j^*}r_{j^*k} - r_{i^*j^*}r_{i^*k} - r_{j^*k}r_{i^*k}) - \big(r'_{i^*k} + r'_{i^*j^*}r'_{j^*k} - r'_{i^*j^*}r'_{i^*k} - r'_{j^*k}r'_{i^*k}\big) \,\big]$$

$$= \sum_{k<i^*} \delta^{(ki^*j^*)}_{\text{trans}} (-r_{ki^*} + r_{kj^*}) + \sum_{i^*<k<j^*} \delta^{(i^*kj^*)}_{\text{trans}} (-1 + r_{i^*k} + r_{kj^*}) + \sum_{j^*<k} \delta^{(i^*j^*k)}_{\text{trans}} (-r_{j^*k} + r_{i^*k})$$

$$\le \sum_{k<i^*} \delta^{(ki^*j^*)}_{\text{trans}} + \sum_{i^*<k<j^*} \delta^{(i^*kj^*)}_{\text{trans}} + \sum_{j^*<k} \delta^{(i^*j^*k)}_{\text{trans}} \le (n-2) \max_{k\notin\{i^*,j^*\}} \delta^{(i^*j^*k)}_{\text{trans}}. \tag{4.52}$$

Together, the above imply

$$\begin{aligned} H_{\text{cycle}}(\mathbf{d},\mathbf{r}) - H_{\text{cycle}}(\mathbf{d},\mathbf{r}') &= H_{\text{consist}}(\mathbf{d},\mathbf{r}) - H_{\text{consist}}(\mathbf{d},\mathbf{r}') + H_{\text{trans}}(\mathbf{r}) - H_{\text{trans}}(\mathbf{r}') \\ &\ge \delta^{(i^*j^*)}_{\text{consist}} - (n-2) \max_{k\notin\{i,j\}} \delta^{(ijk)}_{\text{trans}} > 0. \end{aligned} \tag{4.53}$$

□

**Claim 7** (Sufficiency of "Transitivity" Penalty Weights)**.** *If* $\delta^{(ij)}_{\text{consist}} > (n-2)\max_{k\notin\{i,j\}} \delta^{(ijk)}_{\text{trans}}$ *for* $1 \le i < j \le n$ *and* $\delta^{(ijk)}_{\text{trans}} = \delta_{\text{trans}} > \max_{\substack{1\le i',j'\le n \\ i'\neq j'}} \Delta_{i'j'}$ *for* $1 \le i < j < k \le n$*, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *does not contain a 2-cycle but does contain a directed cycle there is some* $\mathbf{d}'$ *such that* $G(\mathbf{d}')$ *is a DAG and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d},\mathbf{y},\mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}',\mathbf{y},\mathbf{r})$.

*Proof.* Consider an arbitrary $\mathbf{d}^{(l)}$ such that $G(\mathbf{d}^{(l)})$ does not contain a 2-cycle but does contain a directed cycle. Let $\mathbf{r}^{(l)} \equiv \arg\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)},\mathbf{r})$. By Claim 6, $H_{\text{consist}}(\mathbf{d}^{(l)},\mathbf{r}^{(l)}) = 0$ and so $H_{\text{cycle}}(\mathbf{d}^{(l)},\mathbf{r}^{(l)}) = H_{\text{trans}}(\mathbf{d}^{(l)},\mathbf{r}^{(l)})$.

If $\delta^{(ijk)}_{\text{trans}} = \delta_{\text{trans}}$ for $1 \le i < j < k \le n$, i.e. the trasitivity penalty weight is uniform for all directed triangles, then $H_{\text{trans}}(\mathbf{r})$ is equal to the product of $\delta_{\text{trans}}$ and the number of directed triangles in the tournament $G(\mathbf{r})$ for all $\mathbf{r}$.

In any tournament with a positive number of directed triangles, there is always some arc whose switch of direction lowers the number of directed triangles. Let $(i^*, j^*)$ be such such an arc for $\mathbf{r}^{(l)}$. Define $\tilde{\mathbf{r}}^{(l)}$ such that $\tilde{r}_{ij}^{(l)} = \begin{cases} r_{ij}^{(l)}, & (i,j) \neq (i^*, j^*) \\ 1 - r_{ij}^{(l)}, & (i,j) = (i^*, j^*) \end{cases}$. By construction, $H_{\text{trans}}(\mathbf{r}_l^*) - H_{\text{trans}}(\tilde{\mathbf{r}}^{(l)}) \geq \delta_{\text{trans}}$.

It must be the case that $d_{i^*j^*}^{(l)} = 1$. Suppose otherwise. Define some $\mathbf{r}'$ such that $r'_{ij} = \begin{cases} r_{ij}, & (i,j) \neq (i^*, j^*), \\ 1 - r_{ij}, & (i,j) = (i^*, j^*), \end{cases}$ which would have the properties that $H_{\text{consist}}(\mathbf{d}^{(l)}, \mathbf{r}') = 0$ and, by construction, $H_{\text{trans}}(\mathbf{r}') < H_{\text{trans}}(\mathbf{r}^{(l)}$, so that $H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}') < H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}^{(l)}) \neq \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r})$. Because $G(\mathbf{d}^{(l)})$ does not contain a 2-cycle, $d_{j^*i^*}^{(l)} = 0$.

Now, define $\mathbf{d}^{(l+1)}$ such that $d_{ij}^{(l+1)} = \begin{cases} d_{ij}^{(l)}, & (i,j) \neq (i^*, j^*), \\ 0 = 1 - d_{ij}^{(l)}, & (i,j) = (i^*, j^*). \end{cases}$ Then

$$\begin{aligned} H_{\text{consist}}(\mathbf{d}^{(l+1)}, \tilde{\mathbf{r}}^{(l)}) &= \sum_{\substack{1 \leq i < j \leq n \\ (i,j) \neq (i^*, j^*)}} H_{\text{consist}}^{(ij)}(d_{ij}^{(l+1)}, d_{ji}^{(l+1)}, \tilde{r}_{ij}^{(l)}) + H_{\text{consist}}^{(i^*j^*)}(d_{i^*j^*}^{(l+1)}, d_{j^*i^*}^{(l+1)}, \tilde{r}_{i*j*}^{(l)}) \\ &= \sum_{\substack{1 \leq i < j \leq n \\ (i,j) \neq (i^*, j^*)}} H_{\text{consist}}^{(ij)}(d_{ij}^{(l)}, d_{ji}^{(l)}, r_{ij}^{(l)}) + H_{\text{consist}}^{(i^*j^*)}(0, 0, r_{i*j*}^{(l)}) \\ &\leq H_{\text{consist}}(\mathbf{d}^{(l)}, \mathbf{r}^{(l)}) = 0 \end{aligned} \tag{4.54}$$

Because of this, it must be that $H_{\text{trans}}(\tilde{\mathbf{r}}^{(l)}) \geq H_{\text{trans}}(\mathbf{r}^{(l+1)})$, whose negation contradicts the definition of $\mathbf{r}^{(l+1)}$. Therefore,

$$\begin{aligned} H_{\text{trans}}(\mathbf{r}^{(l)}) - H_{\text{trans}}(\mathbf{r}^{(l+1)}) &\geq H_{\text{trans}}(\mathbf{r}^{(l)}) - H_{\text{trans}}(\tilde{\mathbf{r}}^{(l)}) \\ &\geq \delta_{\text{trans}} > \Delta_{i^*, j^*} \\ &\geq H_{\text{score}}(\mathbf{d}^{(l+1)}) - H_{\text{score}}(\mathbf{d}^{(l)}). \end{aligned} \tag{4.55}$$

Because $\mathbf{d}^{(l+1)} \leq \mathbf{d}^{(l)}$, $G(\mathbf{d}^{(l+1)}$ also does not contain a 2-cycle and, by Claim 1,

$$\min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l+1)}, \mathbf{y}) \leq \min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l)}, \mathbf{y}), \tag{4.56}$$

which, together with the above, implies

$$\begin{aligned}
&\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(l)}, \mathbf{y}, \mathbf{r}) - \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(l+1)}, \mathbf{y}, \mathbf{r}) = H_{\text{score}}(\mathbf{d}^{(l)}) - H_{\text{score}}(\mathbf{d}^{(l+1)}) \\
&\quad + \min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l)}, \mathbf{y}) - \min_{\mathbf{y}} H_{\max}(\mathbf{d}^{(l+1)}, \mathbf{y}) + \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l)}, \mathbf{r}) - \min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}^{(l+1)}, \mathbf{r}) \\
&\geq H_{\text{score}}(\mathbf{d}^{(l)}) - H_{\text{score}}(\mathbf{d}^{(l+1)}) + H_{\text{trans}}(\mathbf{r}^{(l)}) - H_{\text{trans}}(\mathbf{r}^{(l+1)}) \qquad > 0.
\end{aligned} \tag{4.57}$$

Let $\mathbf{d}^{(0)} \equiv \mathbf{d}$. Because there can be only finitely many directed triangles in a graph of fixed order, we can construct a sequence $\mathbf{d}, \mathbf{d}^{(1)}, \ldots, \mathbf{d}^{(l)}, \ldots, \mathbf{d}^{(l^*)}$ such that $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}^{(l^*)}, \mathbf{y}, \mathbf{r})$ and $G(\mathbf{r}_{l^*})$ does not contain directed triangle. Thus $H_{\text{trans}}(\mathbf{r}_{l^*}) = H_{\text{cycle}}(\mathbf{d}^{(l^*)}, \mathbf{r}^*_{l^*}) = 0$, which means that $G(\mathbf{d}^{(l^*)})$ is a DAG. Setting $\mathbf{d}' \equiv \mathbf{d}^{(l^*)}$ completes the proof.

□

**Claim 8** (Sufficiency of "Cycle" Penalty Weights)**.** *If* $\delta^{(ij)}_{\text{consist}} > (n-2) \max_{k \notin \{i,j\}} \delta^{(ijk)}_{\text{trans}}$ *for all* $1 \leq i < j \leq n$ *and* $\delta^{(ijk)}_{\text{trans}} = \delta_{\text{trans}} > \max_{\substack{1 \leq i,j \leq n \\ i \neq j}}$ *for all* $1 \leq i < j < k \leq n$*, then for all* $\mathbf{d}$ *such that* $G(\mathbf{d})$ *contains a directed cycle, there is a* $\mathbf{d}' \leq \mathbf{d}$ *such that* $G(\mathbf{d}')$ *is a DAG, and* $\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r}) < \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r})$.

*Proof.* If $G(\mathbf{d})$ contains a 2-cycle, then by Claim 5, there is some $\mathbf{d}'' \leq \mathbf{d}$ such that $G(\mathbf{d}'')$ does not contain a directed 2-cycle and

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}'', \mathbf{y}, \mathbf{r}). \tag{4.58}$$

If $G(\mathbf{d}'')$ is a DAG, then setting $\mathbf{d}' \equiv \mathbf{d}''$ completes the proof. If $G(\mathbf{d})$ does not contain a 2-cycle, set $\mathbf{d}'' \equiv \mathbf{d}$. Then by Claim 7, there is a $\mathbf{d}' \leq \mathbf{d}''$ such that $G(\mathbf{d}')$ is a DAG and

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}'', \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r}). \tag{4.59}$$

□

**Claim 9** (Overall Sufficiency)**.** *If* $\delta_{\max}^{(i)} > \max_{j \neq i} \Delta_{ji}$ *for all* $1 \leq i \leq n$, $\delta_{\text{consist}}^{(ij)} > (n-2) \max_{k \notin \{i,j\}} \delta_{\text{trans}}^{(ijk)}$ *for all* $1 \leq i < j \leq n$ *and* $\delta_{\text{trans}}^{(ijk)} = \delta_{\text{trans}} > \max_{\substack{1 \leq i', j' \leq n \\ i' \neq j'}} \Delta_{i'j'}$ *for all* $1 \leq i < j < k \leq n$, *then* $H(\mathbf{d}^*, \mathbf{y}, \mathbf{r}) = \min_{\substack{\max_i |\mathbf{d}_i| \leq m \\ G(\mathbf{d}) \text{ is a DAG}}} H_{\text{score}}(\mathbf{d}, \mathbf{y}, \mathbf{r})$, $G(\mathbf{d}^*)$ *is a DAG, and* $\max_i |\mathbf{d}_i^*| \leq m$, *where* $\mathbf{d}^* = \arg\min_{\mathbf{d}} \{\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r})\}$.

*Proof.* Consider an arbitrary $\mathbf{d}$. If $\max_i |\mathbf{d}_i| > m$, then by Claim 4, there exists some $\mathbf{d}'$ such that

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}', \mathbf{y}, \mathbf{r}) \tag{4.60}$$

and $\min_{\mathbf{y}} H_{\max}(\mathbf{d}') = 0$, i.e. $\max_i |\mathbf{d}_i'| \leq m$. If $G(\mathbf{d})$ has a directed cycle, then by Claim 8, there is some $\mathbf{d}''$ such that

$$\min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}, \mathbf{y}, \mathbf{r}) > \min_{\mathbf{y},\mathbf{r}} H(\mathbf{d}'', \mathbf{y}, \mathbf{r}) \tag{4.61}$$

and $\min_{\mathbf{r}} H_{\text{cycle}}(\mathbf{d}'', \mathbf{r}) = 0$, i.e. $G(\mathbf{d}'')$ is a DAG. Either of these cases implies that $\mathbf{d} \neq \mathbf{d}^*$, so it must be that $G(\mathbf{d}^*)$ is a DAG, $\max_i |\mathbf{d}_i^*| \leq m$, and $H(\mathbf{d}^*, \mathbf{y}, \mathbf{r}) = \min_{\substack{\max_i |\mathbf{d}_i| \leq m \\ G(\mathbf{d}) \text{ is a DAG}}} H_{\text{score}}(\mathbf{d}, \mathbf{y}, \mathbf{r})$.

□

## Chapter 5

# Training Robust Binary Classifiers Using Quantum Annealing

Apart from minor modifications, this chapter originally appeared as [13]:

"Construction of Non-Convex Polynomial Loss Functions for Training a Binary Classifier with Quantum Annealing". Ryan Babbush, Vasil Denchev, Nan Ding, Sergei Isakov and Hartmut Neven. arXiv preprint 1406.4203. 1-9. 2014.

## Abstract

Quantum annealing is a heuristic quantum algorithm which exploits quantum resources to minimize an objective function embedded as the energy levels of a programmable physical system. To take advantage of a potential quantum advantage for training binary classifiers, one needs to be able to map the regularized risk function to the native hardware with reasonably low overhead. Because experimental considerations constrain our training objective to take the form of a low degree PUBO (POLYNOMIAL UNCONSTRAINED BINARY OPTIMIZATION), we employ non-convex loss functions which are polynomial functions of the margin. We show that these loss functions are robust to label noise and provide a clear advantage over convex methods. These loss functions may also be useful for classical approaches as they compile to regularized risk expressions which can be evaluated in constant time with respect to the number of training examples.

## 5.1 Introduction

### 5.1.1 Quantum annealing

While it is well known that gate model quantum algorithms provide an exponential speedup over the best known classical approaches for some problems [251, 169], we are still technologically far from the ability to construct a large scale quantum computer which can robustly implement such algorithms for nontrivial problem instances. By contrast, rapid advances in superconducting qubit technology [25] have provided a scalable platform for engineering medium-scale, controllable quantum systems at finite temperature. Such devices would be able to implement a quantum version of simulated annealing [167] known as quantum annealing [159, 104, 238, 255].

Because it is NP-Hard to determine the lowest energy configuration of a system of binary spins subject to controllable linear and quadratic energy terms [23], the ability to engineer and cool such a system provides an approach to solving any optimization problem in the class NP. In general, we do not expect that any device can efficiently solve instances of NP-Hard problems in the worst case. However, there is evidence that quantum resources such as tunneling and entanglement are generic computational resources which may help to solve problem instances which would be otherwise intractable for classical solvers. For instance, quantum annealing allows disordered magnets to relax to states of higher magnetic susceptibility asymptotically faster than classical annealing [57] and can solve certain oracular problems exponentially faster than any classical algorithm [256]. For the last few years, *D-Wave Systems* has been commercially manufacturing quantum annealing machines [154]. These machines are the subject of ongoing scientific investigations by several third parties which aim to characterize the extent to which the hardware utilizes quantum resources and

whether a scaling advantage is apparent for any class of problems [43, 44, 233].

### 5.1.2 Training under non-convex loss

The problem we consider in this work is the training of a linear binary classifier using noisy data [40]. We assume that the training data is provided as a matrix $\boldsymbol{x} \in \mathbb{R}^{m \times n}$ with the $m$ rows corresponding to unique descriptor vectors containing $n$ features. We are also provided with a vector of labels, $\boldsymbol{y} \in \{-1, 1\}^m$, which associate a binary classification with each feature vector. The training problem is to determine an optimal classifier $\boldsymbol{w} \in \mathbb{R}^n$ which predicts the data by classifying example $i$ as $\text{sign}\left(\boldsymbol{w}^\top \boldsymbol{x}_i\right)$.

The classifier may be viewed as a hyperplane in feature space which divides data points into negative and positive classifications. In this space, the distance that example $i$ falls from the classification hyperplane $\boldsymbol{w}$ is referred to as the margin $\gamma_i \equiv y_i \, \boldsymbol{x}_i^\top \boldsymbol{w}$. Whereas a negative margin represents a classification opposite the training label, a positive margin represents a classification consistent with the training label. To cast training as an optimization problem we use the concept of a loss function which penalizes the classification of each example according to its margin [40]. Perhaps the simplest loss function is the 0-1 loss function which provides a correct classification with penalty 0 and an incorrect classification with penalty 1,

$$L_{01}\left(\gamma_i\right) \equiv \frac{1 - \text{sign}\left(\gamma_i\right)}{2}. \tag{5.1}$$

The training objective (known in machine learning as total empirical risk) is given as the mean loss over all examples in the training set. For instance, the 0-1 empirical risk function is

$$f_{01}\left(\boldsymbol{w}\right) \equiv \frac{1}{m} \sum_{i=0}^{m-1} L_{01}\left(\gamma_i\right). \tag{5.2}$$

Unfortunately, minimization of the 0-1 empirical risk function is known to be NP-

HARD [106]. For this reason, most contemporary research focuses on convex loss functions which are provably efficient to optimize. However, in data with high label noise, this is an unacceptable compromise as the efficiency gained by convex minimization allows only for the efficient computation of a poor classifier [191]. By contrast, training under non-convex loss functions is known to provide robust classifiers even when nearly half of the examples are mislabeled [184].

Objectives such as these, for which certain instances may require exponential time using classical heuristics, are ideal candidates for quantum annealing. In order to attempt non-convex risk minimization with quantum annealing in the near future, one must first *efficiently* compile the problem to a form compatible with quantum hardware. Due to engineering considerations, this usually means preparing the problem as an instance of QUBO (QUADRATIC UNCONSTRAINED BINARY OPTIMIZATION). Previously, Denchev *at al.* introduced a method for mapping non-convex loss training to QUBO [86] for the purposes of solving on a quantum device. However, in that work, the number of variables required to accomplish the embedding was lower-bounded by the number of training examples. While clearly robust, this scheme seems impractical for medium-scale quantum annealers due to the large qubit overhead.

Here, we develop a different embedding in which the number of required variables is independent of the number of training examples. This is accomplished by deriving loss functions which are polynomial functions of the margins. We show that such loss functions give rise to empirical risk objectives expressible as PUBO. Compatibility with quantum hardware comes from the fact that any PUBO can be reduced to QUBO using a number of boolean ancilla variables that is at most $O\left(N^{2\log k}\right)$ where $N$ is the number of logical variables and $k$ is the order of the PUBO [49]. Coincidentally, this implies that the empirical

risk objective associated with any polynomial loss function can be evaluated in an amount of time that does not depend on the number of training examples.

In particular, we investigate the use of third-order and sixth-order polynomial loss functions. The cubic loss function is chosen as $k = 3$ is the lowest order that gives us non-convexity. Polynomial loss has very different characteristics depending on the parity of $k$ so we also investigate an even degree polynomial loss function. We forgo quartic loss in favor of sixth-order loss as the latter qualitatively fits 0-1 loss much better than the former. After deriving optimal forms of cubic loss and sixth-order loss we numerically investigate the properties of these loss functions to show robustness to label noise. Finally, we demonstrate an explicit mapping of any polynomial risk objective to a tensor representing an instance of PUBO that is easily compiled to quantum hardware.

## 5.2 Cubic loss

In this section we derive an approximate embedding of 0-1 risk under $\ell_2$-norm regularization as a cubic function of the weights. We begin by considering the general forms of the cubic loss and cubic risk functions,

$$L_3(\gamma_i) = \alpha_0 + \alpha_1\gamma_i + \alpha_2\gamma_i^2 + \alpha_3\gamma_i^3 \tag{5.3}$$

$$f_3(\boldsymbol{w}) = \frac{1}{m}\sum_{i=0}^{m-1} L_3(\gamma_i). \tag{5.4}$$

Thus, the embedding problem is to choose the optimal $\boldsymbol{\alpha} \in \mathbb{R}^4$ so that $f_3(\boldsymbol{w})$ best approximates $f_{01}(\boldsymbol{w})$. To accomplish this we consider the $\ell_2$-norm between 0-1 risk and cubic risk,

$$\boldsymbol{\alpha}^* = \operatorname{argmin}\left\{\int P(\boldsymbol{w})\left[f_{01}(\boldsymbol{w}) - f_3(\boldsymbol{w})\right]^2 \mathrm{d}\boldsymbol{w}\right\}. \tag{5.5}$$

Here, $P(\boldsymbol{w})$ is the prior distribution of the weights. If we incorporate an $\ell_2$-norm regularizer, $\Omega_2(\boldsymbol{w})$, into our ultimate training objective, $E(\boldsymbol{w})$, i.e.

$$\Omega_2(\boldsymbol{w}) = \frac{\lambda_2}{2}\boldsymbol{w}^\top\boldsymbol{w} \tag{5.6}$$

$$E(\boldsymbol{w}) = f(\boldsymbol{w}) + \Omega_2(\boldsymbol{w}), \tag{5.7}$$

then we are provided with a Gaussian prior on the weights [227] taking the form,

$$P(w_i) = \sqrt{\frac{\lambda_2}{2\pi}}e^{-\lambda_2 w_i^2/2}. \tag{5.8}$$

Immediately, we see that for the optimal solution $\alpha_2 \to 0$ since 0-1 loss is an odd function and the least squares residual is weighed over a symmetric function (the Gaussian prior). Furthermore, we can ignore $\alpha_0$ and the constant factor of $\frac{1}{2}$ in $L_{01}(\gamma_i)$ as these constants are irrelevant for the training problem. With this in mind, we expand the empirical risk functions under the integral in the embedding problem as,

$$\int P(\boldsymbol{w})\left(\sum_{i=0}^{m-1}\frac{\text{sign}(\gamma_i)}{2} + \alpha_1\gamma_i + \alpha_3\gamma_i^3\right)^2 \mathrm{d}\boldsymbol{w}. \tag{5.9}$$

Thus,

$$\boldsymbol{\alpha}^* = \text{argmin}\left\{\sum_{i=0}^{m-1}\sum_{j=0}^{m-1}\int P(\boldsymbol{\gamma})F_{ij}(\boldsymbol{\gamma})\,\mathrm{d}\boldsymbol{\gamma}\right\} \tag{5.10}$$

where

$$\begin{aligned} F_{ij}(\boldsymbol{\gamma}) \equiv\ & \frac{\alpha_1}{2}\left[\gamma_i\text{sign}(\gamma_j) + \gamma_j\text{sign}(\gamma_i)\right] + \frac{\alpha_3}{2}\left[\gamma_i^3\text{sign}(\gamma_j) + \gamma_j^3\text{sign}(\gamma_i)\right] \\ & + \alpha_1^2\gamma_i\gamma_j + \alpha_1\alpha_3\left(\gamma_i^3\gamma_j + \gamma_j^3\gamma_i\right) + \alpha_3^2\gamma_i^3\gamma_j^3. \end{aligned} \tag{5.11}$$

Without loss of generality, we may assume that $P(\boldsymbol{\gamma})$ is a multinormal distribution centered at zero with a covariance matrix,

$$\boldsymbol{\Sigma} = \frac{1}{\lambda_2}\left(\boldsymbol{x}^\top\boldsymbol{x}\right) \odot \left(\boldsymbol{y}\,\boldsymbol{y}^\top\right) \tag{5.12}$$

where $\odot$ implies element-wise matrix multiplication (i.e. the Hadamard product). The multinormal distribution occurs because the margins arise as the result of the training examples being projected by classifiers drawn from the prior distribution given by $P(\boldsymbol{w})$. Since each weight is normally distributed with zero mean and variance $\lambda_2^{-2}$ in the prior, the distribution of margins associated with training example $i$ will be a Gaussian with zero mean and variance,

$$\sigma_i^2 = \frac{1}{n}\sum_{j=0}^{n-1}\left(\frac{x_{ij}}{\lambda_2}\right)^2. \tag{5.13}$$

Because each linear combination of the elements of the margin vector is also normally distributed, we have a multinormal distribution. This is true regardless of the number of features or any particular qualities of the training data.

Accordingly, if we wished to scale $\boldsymbol{w}$ to a range which contains $\boldsymbol{w}^*$ with a likelihood in the $r^{\text{th}}$ standard deviation of the prior then we should make $\boldsymbol{w} \in \left[-\frac{r}{\sqrt{\lambda_2 m}}, \frac{r}{\sqrt{\lambda_2 m}}\right]^n$. However, making $r$ too large would be problematic because this could allow the cubic term to dominate the quadratic regularizer. This necessitates a cutoff on the maximum weight value to ensure that unbounded cubic losses associated with large negative margins do not overcome the regularizer. In practice, $r$ would need to be selected as a hyperparameter.

Since the integrand has only two point correlation functions, we can integrate over the marginal distribution of $\gamma_i$ and $\gamma_j$ which is a binormal distribution with covariance,

$$\boldsymbol{\Sigma}_{ij} = \frac{1}{\lambda_2}\begin{pmatrix} \boldsymbol{x}_i^\top \boldsymbol{x}_i & y_i y_j \boldsymbol{x}_i^\top \boldsymbol{x}_j \\ y_j y_i \boldsymbol{x}_j^\top \boldsymbol{x}_i & \boldsymbol{x}_j^\top \boldsymbol{x}_j \end{pmatrix} = \begin{pmatrix} \sigma_i^2 & \rho_{ij}\,\sigma_i\sigma_j \\ \rho_{ij}\,\sigma_j\sigma_i & \sigma_j^2 \end{pmatrix}. \tag{5.14}$$

We can analytically evaluate the double integral,

$$\boldsymbol{\alpha}^* = \operatorname{argmin} \left\{ \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} I_{ij} \right\} \tag{5.15}$$

$$\begin{aligned} I_{ij} &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} P_{ij}\left(\boldsymbol{\gamma}\right) F_{ij}\left(\boldsymbol{\gamma}\right) \mathrm{d}\gamma_i \mathrm{d}\gamma_j \\ &= \underbrace{\frac{\rho_{ij}\left(\sigma_i + \sigma_j\right)}{\sqrt{2\pi}}}_{t_0} \alpha_1 + \underbrace{\frac{\rho_{ij}\left(3 - \rho_{ij}^2\right)\left(\sigma_i^3 + \sigma_j^3\right)}{\sqrt{2\pi}}}_{t_1} \alpha_3 \\ &\quad + \underbrace{3\,\rho_{ij}\sigma_i\sigma_j\left(\sigma_i^2 + \sigma_j^2\right)}_{t_2} \alpha_1\alpha_3 + \underbrace{\rho_{ij}\sigma_i\sigma_j}_{t_3} \alpha_1^2 + \underbrace{3\,\rho_{ij}\left(3 + 2\rho_{ij}^2\right)\sigma_i^3\sigma_j^3}_{t_4} \alpha_3^2 \\ &= t_0\alpha_1 + t_1\alpha_3 + t_2\alpha_1\alpha_3 + t_3\alpha_1^2 + t_4\alpha_3^2 \end{aligned} \tag{5.16}$$

With the integral in closed form, we obtain the argmin using simple algebra by solving,

$$\nabla \left( \sum_{i=0}^{m-1} \sum_{j=0}^{m-1} I_{ij} \right) = 0. \tag{5.17}$$

The analytical coefficients for a single set of training examples are,

$$\alpha_1^* = \frac{2\,t_0t_4 - t_1t_2}{t_2^2 - 4\,t_3t_4} \qquad \alpha_3^* = \frac{2\,t_1t_3 - t_0t_2}{t_2^2 - 4\,t_3t_4}. \tag{5.18}$$

Thus, for the full training set we must sum together the $t$ values from each set of $(i, j)$ before plugging values into the expression for $\alpha_1$ and $\alpha_3$. We note that the computation required to obtain these coefficients is $O\left(m^2\right)$. Figure 5.1 shows several cubic loss function fits associated with various real data sets from the UCI Machine Learning Repository.

The coefficients above are analytic and optimal for embedding 0-1 risk. However, it is instructive to explain what would happen if we had chosen to fit the loss function instead of the objective function. This would have produced a far simpler embedding problem[1],

$$\boldsymbol{\alpha}^\star = \operatorname{argmin} \left\{ \int_{-\infty}^{\infty} P^\star\left(\gamma\right) \left[L_{01}\left(\gamma\right) - L_3\left(\gamma\right)\right]^2 \mathrm{d}\gamma \right\} \tag{5.19}$$

[1]Note the difference between $\boldsymbol{\alpha}^\star$ and $\boldsymbol{\alpha}^*$.

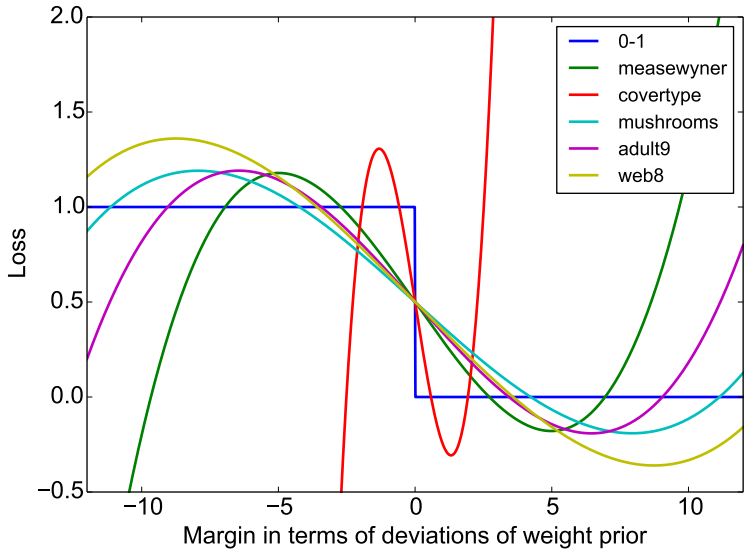



Figure 5.1: Cubic loss fits for a variety of real data sets. Due to the different properties of their correlation matrices each set is associated with unique cubic loss coefficients.

where

$$P^{\star}(\gamma) \equiv \frac{1}{\sigma\sqrt{2\pi}} e^{-\gamma^2/2\sigma^2}, \; \sigma^2 = \frac{1}{\lambda_2 m} \mathrm{tr}\left[\boldsymbol{x}^{\top}\boldsymbol{x}\right]. \tag{5.20}$$

This time the integral is trivial to evaluate,

$$\begin{aligned} I^{\star} &= \int_{-\infty}^{\infty} G^{\star}(\gamma,) \left[L_{01}(\gamma) - L_3(\gamma)\right]^2 \mathrm{d}\gamma \\ &= \sqrt{\frac{2}{\pi}}\sigma\alpha_1 + \sigma^2\alpha_1^2 + 2\sqrt{\frac{2}{\pi}}\sigma^3\alpha_3 \\ &\quad + 6\,\sigma^4\alpha_1\alpha_3 + 15\,\sigma^6\alpha_3^2. \end{aligned} \tag{5.21}$$

As before, convexity guarantees that $\nabla I$ will have exactly one real root which we find analytically,

$$\alpha_1^{\star} = -\frac{3}{2\sqrt{2\pi}\sigma} \qquad \alpha_3^{\star} = \frac{1}{6\sqrt{2\pi}\sigma^3}. \tag{5.22}$$

These result seems much simpler than when we embed the entire risk function but they are obviously less useful. However, if we make the assumption that $\sigma_i = \sigma_j = \sigma$ then, $\alpha_1^{\star} = \alpha_1^{*}$ and $\alpha_3^{\star} = \alpha_3^{*}$.

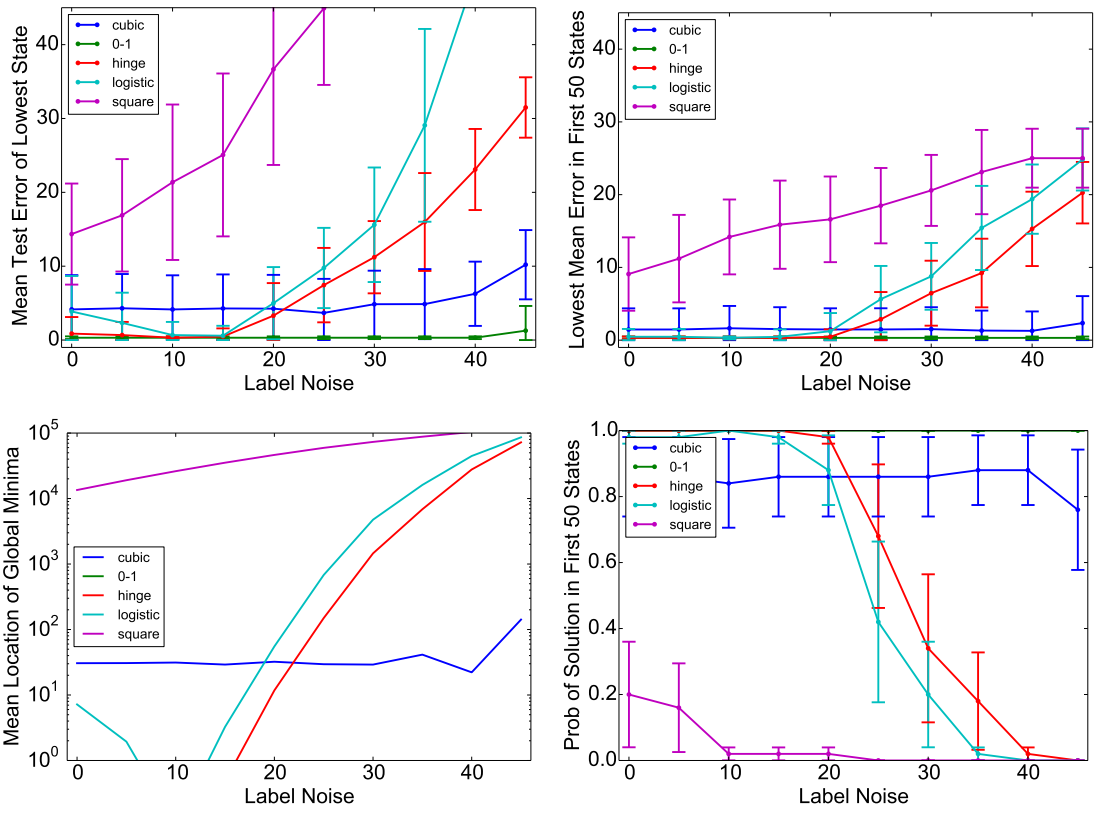



Figure 5.2: Performance of cubic loss on synthetic data sets with $10^4$ examples, 9 features and a bit depth of 2. We exactly enumerate all fixed bit depth classifiers and evaluate the empirical risk under various loss functions. Error bars are obtained by repeating the experiment on fifty data sets. The upper left plot shows the error in the lowest objective state and the upper right plot shows the error in the best state of the lowest fifty. The lower left plot shows mean position of the global minima in the eigenspectra of the various objective functions. The lower right plot shows the probability of the global minima being in the first fifty states. As we can see, the global minima remains very near the bottom of the eigenspectra for cubic loss, regardless of label noise.

In Figure 5.2 we study the performance of our embedded loss function by exactly enumerating the solution space produced by small synthetic data sets. These data sets were produced by randomly generating classifiers with weights drawn from a normal distribution and then using that classifier to label feature vectors with features drawn from a uniform distribution. Symmetric label errors are then manually injected at random. A maximum weight cutoff is imposed at the second standard deviation of the weight prior. Further numerical analysis of cubic loss on synthetic data sets is included in the Appendix.

While the cubic loss embedding is somewhat noisy in the sense that it does not perfectly approximate 0-1 loss, it is clearly robust in the sense that test error does not depend strongly on label error for up to 45% label noise. This remains true whether we consider the best fifty states embedded in cubic loss or only the absolute ground state. These results indicate that cubic loss has an advantage over convex methods when data is known to contain substantial label noise.

## 5.3 Sixth-order loss

One potentially unattractive feature of the cubic loss function is that it is necessary to fix the scale of the weights as a hyperparameter. Since we intend to encode our objective function as QUBO for quantum annealing, we will need to choose a maximum weight. While one can prove that the optimal classifier will have weights in the interval $\left[-\frac{1}{\sqrt{\lambda_2}}, \frac{1}{\sqrt{\lambda_2}}\right]$, such a large range is potentially problematic for regularized cubic risk as the loss associated with large negative margins tends towards negative infinity faster than the $\ell_2$ regularizer can penalize the large weights which would produce such margins.

Alternatively, one might consider using a polynomial loss function of even degree as such loss functions will not diverge to negative infinity for large negative margins. In order

to do this, we must fix the highest order term in the loss function as a hyperparameter. This is because 0-1 loss is an odd function so if we attempt to embed 0-1 risk in an even degree polynomial, the even terms will vanish. Accordingly, we turn our attention to the sixth-order loss and empirical risk functions,

$$L_6(\gamma_i) \equiv \omega\gamma_i^6 + \sum_{k=1}^{5} \beta_k \gamma_i^k \tag{5.23}$$

$$f_6(\boldsymbol{w}) \equiv \frac{1}{m} \sum_{i=0}^{m-1} L_3(\gamma_i). \tag{5.24}$$

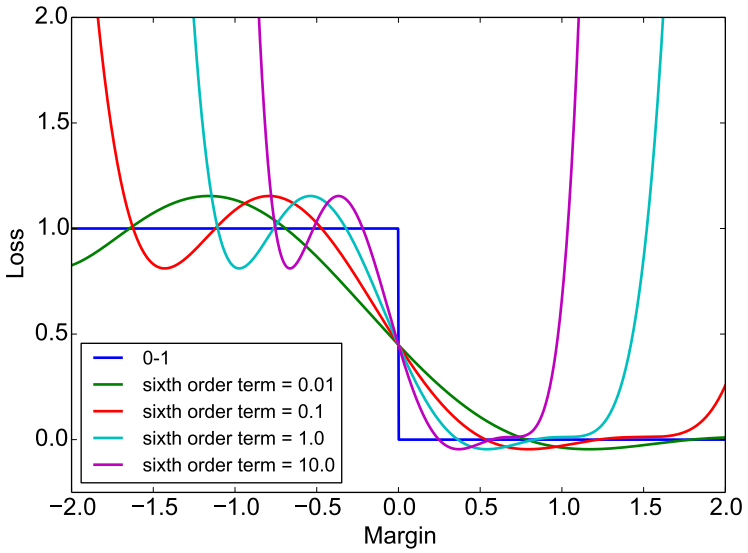



Figure 5.3: The sixth-order loss function at various values of the fixed sixth-order coefficient, $\omega$. This coefficient is taken to be a hyperparameter.

Here, $\omega$ is taken to be a hyperparameter. We will solve for the values of $\boldsymbol{\beta}$. In the case of the cubic loss function, we used the weight prior imposed by $\ell_2$-norm regularization and data set covariance to derive a margin prior which was used for embedding. However, this is unnecessary for the sixth-order loss function as $\omega$ provides a very simple prior on the margins,

$$P(\gamma) = \frac{\omega^{1/6}}{2\,\Gamma(7/6)} e^{-\omega\gamma^6} \tag{5.25}$$

where $\Gamma$ is the standard gamma function. With this definition, the embedding problem becomes

$$\boldsymbol{\beta}^* = \operatorname{argmin} \left\{ \int P(\gamma) \left[ L_{01}(\gamma) - L_6(\gamma) \right]^2 \mathrm{d}\gamma \right\}. \tag{5.26}$$

Whereas we chose $\boldsymbol{\alpha}$ for cubic loss by fitting the empirical risk function, we choose $\boldsymbol{\beta}$ for sixth-order loss by fitting the loss function directly. Since the sixth-order loss function is already parameterized in terms of a hyperparameter, there is nothing to gain by devising a more elaborate fit based on empirical risk. After evaluating $I_6$, the integral in Eq. 5.26, we can obtain $\boldsymbol{\beta}^*$ by solving, $\nabla I_6 = 0$. The optimal values of $\boldsymbol{\beta}$ are included in the Appendix. Figure 5.3 shows the sixth-order loss function for various values of $\omega$.

Figure 5.4 shows the performance of the sixth-order loss function on selected data sets from the UCI Machine Learning Repository. To stand-in for a quantum annealer, we optimized the sixth-order objective function using a simulated annealing routine which was run for over one hundred thousand CPU hours. In addition to standard convex methods, we compare to 0-1 loss optimized using the same simulated annealing code run with the same number of restarts and variable updates as were used to optimize sixth-order loss. We also include the "q-loss" results from [86] which were obtained for that work using another metaheuristic algorithm (Tabu search). Details regarding the 10-fold cross validation procedure are reported in the Appendix.

We find that for all real data sets, the sixth-order loss function outperforms all tested convex loss functions and performs similarly to the non-convex methods. The first two data sets shown in Figure 5.4 are synthetic sets designed to break convex loss functions devised by Long and Servedio [184] and Mease and Wyner [196]. The sixth-order loss performs poorly on the Long-Servedio set because the data set is designed so that the optimal solution has only extremely large margins and extremely small margins. The large margins dominate

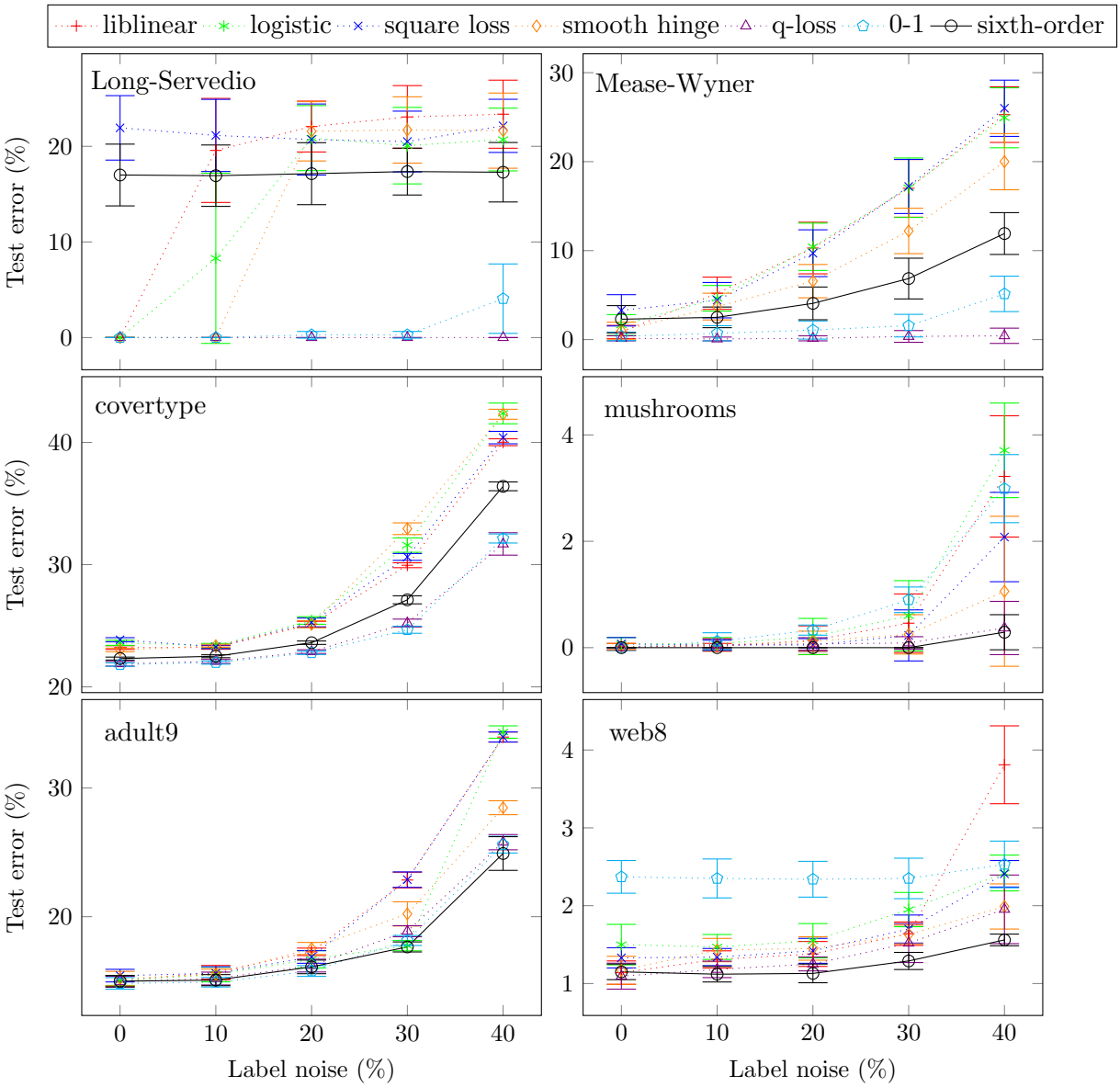



Figure 5.4: Test error versus label noise for 7 methods on 2 synthetic data sets (Long-Servedio and Mease-Wyner) and 4 real data sets from the UCI repository. Error bars are obtained from 10-fold cross-validation with the hyperparameters recorded in the Appendix. As a stand-in for quantum annealing, a classical simulated annealing routine was used to optimize the sixth-order objective function and the 0-1 objective function. For each training cut at each selection of hyperparameters, we kept 50 classifiers having the lowest objective values of all states encountered. We computed validation error as the lowest of validation error produced by these 50 classifiers. This procedure is used for both 0-1 loss and sixth-order loss. Test error was obtained using the classifier of lowest validation error. This strategy is realistic as we expect that a quantum annealer will sample the lowest energy states as opposed to giving us only the global minima. q-loss was optimized using Tabu search in [86]. We see that sixth-order loss outperforms the convex methods on every data set except for Long-Servedio and performs similarly to other non-convex methods on the other five sets.

the risk minimization due to the steep walls of the sixth-order function and this forces all of the smaller margins very near zero where the sixth-order function is almost linear. We believe that the particular pathological behavior which leads to the poor performance of sixth-order loss on the Long-Servedio set is unlikely to occur in real data.

Figure 5.4 shows that the sixth-order loss function outperforms even the other non-convex methods on three of the four real data sets. However, we attribute the suboptimal q-loss and 0-1 loss results to a failure of the selected optimization routines rather than to a deficiency of the actual training objectives. One reason this seems likely is because sixth-order loss outperforms the other non-convex functions most significantly on web8 (the real data set with the greatest number of features) but losses to q-loss and 0-1 loss on covertype (the real data set with the fewest number of features). A comprehensive summary of the data sets is included in the Appendix. While non-convex, the sixth-order loss objective appears to be somewhat easier to optimize as a consequence of being significantly smoother than either the 0-1 loss or q-loss objective.

## 5.4 Explicit tensor construction

In this section we show how to represent any regularized risk objective using a polynomial loss function as PUBO. We first introduce a fixed bit-depth approximation. More substantially, we represent variables using a fixed-point representation as floating-point representations require a non-polynomial function of the bits. Using $d$ bits per feature, our encoding will require a total of $N = nd$ bits. The binary state vector $\boldsymbol{q} \in \mathbb{B}^N$ encodes the weight vector $\boldsymbol{w}$,

$$\boldsymbol{w}[i] \equiv \zeta \boldsymbol{q}[id] - \zeta \sum_{j=1}^{d-1} \left(\frac{1}{2}\right)^j \boldsymbol{q}[id+j] \tag{5.27}$$

where $\zeta \in \mathbb{R}$ determines the weight scale so that $\boldsymbol{w} \in (-\zeta, \zeta]^n$. Furthermore, we define a binary coefficient matrix, $\boldsymbol{k} \in \mathbb{R}^{n \times N}$,

$$\boldsymbol{k} \equiv \boldsymbol{I}^{n \times n} \otimes \left\langle \zeta, -\frac{\zeta}{2}, -\frac{\zeta}{4}, \ldots, \frac{\zeta}{2^{1-d}} \right\rangle \tag{5.28}$$

where $\boldsymbol{I}^{n \times n}\otimes$ indicates a Kronecker product by an $n \times n$ identity matrix. This "tiles" the binary weight sequence into a stair-step pattern down the diagonal; e.g. if $n = 3$, $d = 2$ and $\zeta = 1$,

$$\boldsymbol{k} = \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -\frac{1}{2} \end{pmatrix}. \tag{5.29}$$

We do this because later on, it will be useful to think of $\boldsymbol{w}$ as a linear mapping of $\boldsymbol{q}$ given as $\boldsymbol{w} = \boldsymbol{k}\boldsymbol{q}$. In general, any PUBO of degree $k$ can be expressed as,

$$E\left(\boldsymbol{q}\right) = \boldsymbol{v}^\top \boldsymbol{q}^{\otimes k} \tag{5.30}$$

where $\boldsymbol{v} \in \mathbb{R}^{N^k}$ is a $k$-fold tensor and $\boldsymbol{q}^{\otimes k}$ represents the $k^{\text{th}}$ tensor power of $\boldsymbol{q}$. We now show how to obtain this embedding for a cubic loss function but do so in a way that is trivially extended to orders less than or greater than cubic. In terms of continuous weights the empirical risk objective may be expressed as,

$$f\left(\boldsymbol{w}\right) = \frac{1}{m}\sum_{i=0}^{m-1} \alpha_1 y_i \boldsymbol{x}_i^\top \boldsymbol{w} + \alpha_2 \left(\boldsymbol{x}_i^\top \boldsymbol{w}\right)^2 + \alpha_3 \left(y_i \boldsymbol{x}_i^\top \boldsymbol{w}\right)^3 \tag{5.31}$$

$$= \underbrace{\left(\frac{\alpha_1}{m}\sum_{i=0}^{m-1} y_i \boldsymbol{x}_i^\top\right)}_{\boldsymbol{\varphi_1}^\top} \boldsymbol{w} + \underbrace{\left(\frac{\alpha_2}{m}\sum_{i=0}^{m-1} \left(\boldsymbol{x}_i^{\otimes 2}\right)^\top\right)}_{\boldsymbol{\varphi_2}^\top} \boldsymbol{w}^{\otimes 2} + \underbrace{\left(\frac{\alpha_3}{m}\sum_{i=0}^{m-1} \left(y_i \boldsymbol{x}_i^{\otimes 3}\right)^\top\right)}_{\boldsymbol{\varphi_3}^\top} \boldsymbol{w}^{\otimes 3} = \sum_{j=1}^{3} \boldsymbol{\varphi_j}^\top \boldsymbol{w}^{\otimes j}$$

where

$$\boldsymbol{\varphi}_j = \frac{\alpha_j}{m}\sum_{i=0}^{m-1} \left(y_i \boldsymbol{x}_i\right)^{\otimes j}. \tag{5.32}$$

Using tensor notation, $\ell_2$-norm regularization is

$$\Omega_2(\boldsymbol{w}) = \frac{\lambda_2}{2}\left(\mathbf{1}^{n^2}\right)^\top \boldsymbol{w}^{\otimes 2}, \tag{5.33}$$

where $\mathbf{1}^{n^2}$ denotes a vector of all ones with length $n^2$. We now use $\boldsymbol{k}$ to expand the binary variable tensor,

$$E(\boldsymbol{q}) = f(\boldsymbol{w}) + \Omega_2(\boldsymbol{w}) = \frac{\lambda_2}{2}\left(\mathbf{1}^{n^2}\right)^\top (\boldsymbol{kq})^{\otimes 2} + \sum_{j=1}^{3} \boldsymbol{\varphi_j}^\top (\boldsymbol{kq})^{\otimes j}. \tag{5.34}$$

This expression implies the form of $\boldsymbol{v}$,

$$\boldsymbol{v} = \frac{\lambda_2}{2}\mathbf{1}^{n^2} \otimes \boldsymbol{k}^{\otimes 2} + \sum_{j=1}^{3} \boldsymbol{\varphi_j} \otimes \boldsymbol{k}^{\otimes j}. \tag{5.35}$$

We slightly abuse notation in our definition of $\boldsymbol{v}$ by "adding" together tensors of different rank. To accomplish this the tensor of lower rank should be placed in a tensor having the same rank as the larger tensor by setting additional tensor indices equal to a lower tensor index. For instance, the element corresponding to $(i, j)$ in a tensor of rank two could be placed in a tensor of rank three at $(i, j, i)$ or $(i, j, j)$. We note that it is necessary to *first* convert to binary and then combine the three terms; doing things the other way would introduce complications due to the fact that $w_i^r \neq w_i \ \forall i, r$ whereas $q_i^r = q_i \ \forall i, r$. Finally, we note that constructing $\boldsymbol{\varphi_3}^\top \otimes \boldsymbol{k}^{\otimes 3}$ is the most computationally expensive part of this entire procedure taking $O\left(n^3 d^3 m\right)$ time.

This 3-fold tensor can be reduced to a QUBO matrix using ancillae. The optimal reduction is trivial using the tools developed in [16]. In Appendix B of that paper, it shown that the number of ancillae which are required to collapse a fully connected cubic to 2-local is upper bounded by $\frac{N^2}{4}$. Again, the general bound for the quadratization of a PUBO of degree $k$ is $O\left(N^{2\log k}\right)$ [49]. This bound suggests that unlike prior encodings, the number of ancillae required is entirely independent of the number of training examples. We point

out that the tensor form of this problem may be evaluated in a time that does not depend on the number of training examples.

## 5.5 Conclusion

We have introduced two unusual loss functions: the cubic loss function and the sixth-order loss function. Both losses are non-convex and show clear evidence of robustness to label noise. While superior to classically tractable convex training methods, both loss functions are highly parameterized and represent less than perfect approximations to 0-1 loss. Prima facie, this suggests that more popular non-convex loss functions, e.g. sigmoid loss, may be more reliable (or at least more straightforward) in some respects.

However, training under non-convex loss is formally NP-Hard and in order to obtain satisfactory solutions to such optimization problems, heuristic algorithms must query the objective function many times. Often, the quality of the eventual solution depends on the number of queries the heuristic is allowed. The fact that the polynomial loss functions may be compiled so that each query to the objective is independent of the number of training examples suggests that these loss functions may be more compatible with heuristic optimization routines. This same property ensures that these loss functions can be compiled to a Hamiltonian suitable for quantum annealing using a reasonable number of qubits (estimates of resources requirements for various example problems are included in the Appendix). This efficient embedding in quantum hardware makes binary classification under non-convex polynomial loss a promising target problem to accelerate using a quantum annealing machine.

## 5.6 Appendix

### 5.6.1 $\ell_0$-norm regularization

While $\ell_0$ programming is well known to be NP-Hard, we believe that quantum annealing may allow us to obtain satisfactory minima in many instances. $\ell_0$-norm regularization is often used to train classifiers that are particularly efficient in terms of the number of features required for classification. For the situation in which we would like to train a classifier with binary weights, the regularization function is trivial,

$$\Omega_0(\boldsymbol{q}) = \lambda_0 \sum_{i=0}^{n-1} q_i. \tag{5.36}$$

For multiple bit depth weights, $\ell_0$-norm regularization will require a modest number of ancilla qubits (one for each feature). Using our notation, the regularizer is

$$\Omega_0(\boldsymbol{q}) = \sum_{j=0}^{n-1} \left( \lambda_0 \boldsymbol{q}[N+j] + \phi\left(1 - \boldsymbol{q}[N+j]\right) \sum_{k=0}^{d-1} \boldsymbol{q}[jd+k] \right). \tag{5.37}$$

Minimizing $\Omega(\boldsymbol{q})$ causes the ancillae $\boldsymbol{q}[N+j]$ to act as indicator bits, each of which is 1 if and only if $w_j \neq 0$ and is 0 otherwise. This is achieved by summing the binary variables that take part in a weight variable. Correctness comes from the fact that the binary representation of $w_j = 0$ is when all bits corresponding to that weight are 0. Thus, if even a single bit from weight $w_i$ is on, the objective will either incur a penalty of $\phi$ or will set the ancilla to 0 so as to obtain a penalty of $\lambda_0$ instead. Thus, as long as $\phi$ is sufficiently larger than $\lambda_0$, this function enforces $\ell_0$-norm regularization with weight of $\lambda_0$. This regularizer may be combined with the empirical risk function described previously.

### 5.6.2 Sixth-order loss coefficients

$$
\begin{aligned}
\beta_0 &= \frac{1}{3}+\frac{343\left(125\pi^{3/2}-864\,\Gamma\left(\frac{11}{6}\right)^3\right)}{1296\left(343\,\Gamma\left(\frac{11}{6}\right)^3+750\,\Gamma\left(\frac{13}{6}\right)^3\right)-300125\pi^{3/2}}\\
\beta_1 &= -\frac{35\sqrt[6]{\omega}\left(245\left(1+3\ 2^{2/3}\right)\pi\,\Gamma\left(\frac{4}{3}\right)+36\left(49\,\Gamma\left(\frac{11}{6}\right)\left(\sqrt{\pi}-3\,\Gamma\left(\frac{5}{3}\right)\ \Gamma\left(\frac{11}{6}\right)\right)-60\,\Gamma\left(\frac{13}{6}\right)^2\right)\right)}{3\left(-\frac{222950\pi^{3/2}}{9}+98784\,\Gamma\left(\frac{11}{6}\right)^3+43200\,\Gamma\left(\frac{13}{6}\right)^3\right)}\\
\beta_2 &= \frac{2940\sqrt{\pi}\sqrt[3]{\omega}\left(25\sqrt{\pi}\,\Gamma\left(\frac{13}{6}\right)-42\,\Gamma\left(\frac{11}{6}\right)^2\right)}{300125\pi^{3/2}-1296\left(343\,\Gamma\left(\frac{11}{6}\right)^3+750\,\Gamma\left(\frac{13}{6}\right)^3\right)}\\
\beta_3 &= \frac{3675\sqrt{\pi}\sqrt{\omega}\left(7\sqrt{\pi}+63\left(\sqrt[3]{2}-1\right)\ \Gamma\left(\frac{5}{3}\right)\ \Gamma\left(\frac{11}{6}\right)+18\left(2^{2/3}-3\right)\ \Gamma\left(\frac{4}{3}\right)\ \Gamma\left(\frac{13}{6}\right)\right)}{111475\pi^{3/2}-1296\left(343\,\Gamma\left(\frac{11}{6}\right)^3+150\,\Gamma\left(\frac{13}{6}\right)^3\right)}\\
\beta_4 &= \frac{2100\sqrt{\pi}\omega^{2/3}\left(49\sqrt{\pi}\,\Gamma\left(\frac{11}{6}\right)-60\,\Gamma\left(\frac{13}{6}\right)^2\right)}{432\left(343\,\Gamma\left(\frac{11}{6}\right)^3+750\,\Gamma\left(\frac{13}{6}\right)^3\right)-\frac{300125\pi^{3/2}}{3}}\\
\beta_5 &= \frac{7\omega^{5/6}\left(-7056\,\Gamma\left(\frac{11}{6}\right)^2+1225\left(3\sqrt[3]{2}-1\right)\pi\,\Gamma\left(\frac{5}{3}\right)+600\,\Gamma\left(\frac{13}{6}\right)\left(7\sqrt{\pi}-18\,\Gamma\left(\frac{4}{3}\right)\ \Gamma\left(\frac{13}{6}\right)\right)\right)}{-\frac{222950\pi^{3/2}}{9}+98784\,\Gamma\left(\frac{11}{6}\right)^3+43200\,\Gamma\left(\frac{13}{6}\right)^3}
\end{aligned}
$$

### 5.6.3 Convergence of cubic loss function

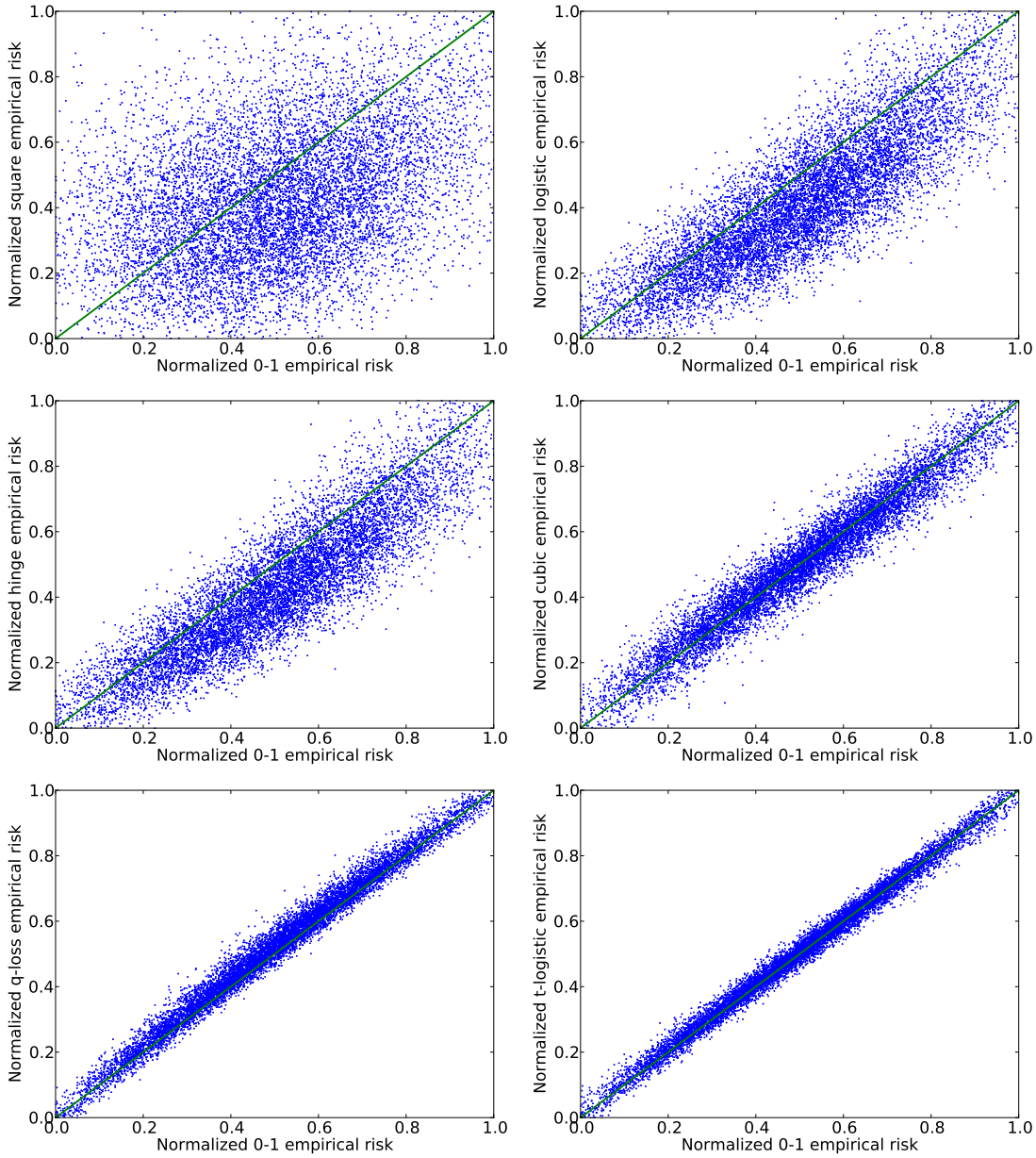



Figure 5.5: Correlations between the total empirical risk of $10^4$ randomly selected states using various loss functions over $10^4$ training examples from the adult9 data set. The empirical risk values of each state have been uniformly shifted and rescaled to be in between 0 and 1. As we can see, the correlations between the convex loss functions and 0-1 loss are strictly worse than the correlation between cubic loss and 0-1 loss.

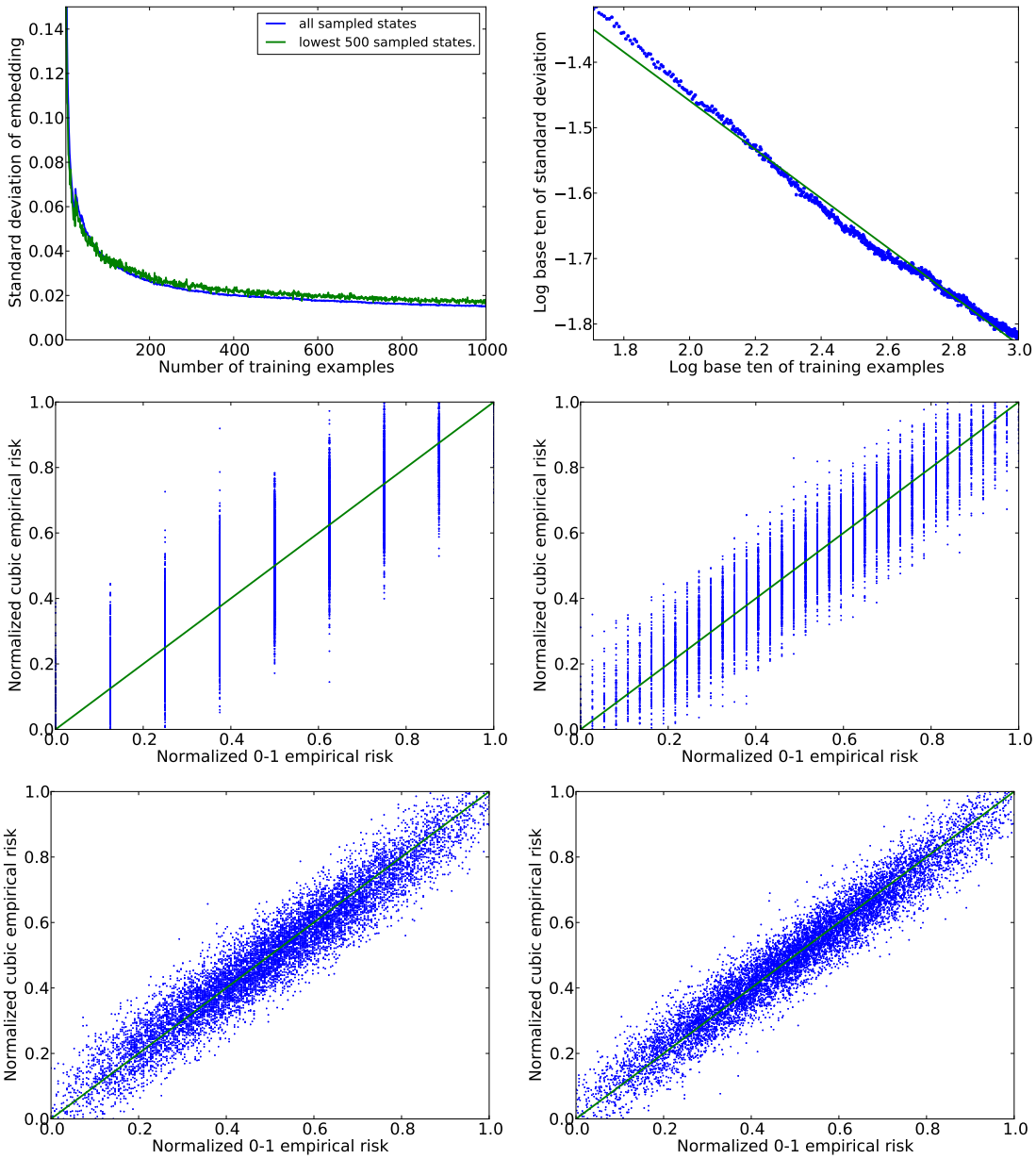



Figure 5.6: The upper-left plot is made by fitting the cubic loss function and evaluating the resultant embedding error on adult9 using a variable number of training examples. Here, the error is the standard deviation of the energy landscapes defined by the limited number of training examples. At each point, $10^4$ states are sampled at random to evaluate the error. We also show the error in only the lowest 50 of these states to give an indication of the rate at which the low energy subspace is converging. On the upper-right, is a log plot of the same data indicating that embedding appears to converge as roughly $O\left(m^{-1/3}\right)$. The remaining plots show correlations between the total empirical risk of the $10^4$ randomly selected states using 0-1 loss and the total empirical risk on those states using cubic loss. These four plots were obtained by fitting the cubic loss function to the adult9 data set using: 10 training examples, 100 training examples, 1,000 training examples and 10,000 training examples. The error in these embeddings are 0.107, 0.0818, 0.062 and 0.055, respectively.

### 5.6.4 Estimated qubit requirements

Table 5.1: Upper-bounds on qubit requirements for selected problems

| Loss function degree | #Features | Bit-depth | #Qubits |
|---|---|---|---|
| cubic | 100 | 1 | 2,550 |
| cubic | 100 | 4 | 40,200 |
| cubic | 500 | 1 | 62,750 |
| cubic | 500 | 4 | 1,001,000 |
| cubic | 2,500 | 1 | 1,563,750 |
| cubic | 2,500 | 4 | 25,005,000 |

### 5.6.5 Data summary

Table 5.2: Data summary

| Name | Dims | #Examples | Density (%) | Baseline error (%) |
|---|---|---|---|---|
| Long-Servedio | 21 | 2000 | 100.00 | 50.00 |
| Mease-Wyner | 20 | 2000 | 100.00 | 49.80 |
| covertype | 54 | 581012 | 22.20 | 36.46 |
| mushrooms | 112 | 8124 | 18.75 | 48.20 |
| adult9 | 123 | 48842 | 11.30 | 23.93 |
| web8 | 300 | 59245 | 4.20 | 2.92 |

### 5.6.6 Hyperparameters

Table 5.3: $\omega$ values for sixth-order loss picked by cross-validation

| Data set name | Label noise (%) | | | | |
|---|---|---|---|---|---|
| | 0 | 10 | 20 | 30 | 40 |
| Long-Servedio | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.000632 |
| Mease-Wyner | 0.398965 | 0.000126 | 0.000126 | 0.000025 | 0.000005 |
| covertype | 2.000000 | 2.000000 | 0.398965 | 2.000000 | 0.000025 |
| mushrooms | 2.000000 | 2.000000 | 2.000000 | 0.000632 | 0.000005 |
| adult9 | 0.000001 | 2.000000 | 0.000025 | 0.079583 | 2.000000 |
| web8 | 0.000632 | 0.000126 | 0.398965 | 0.398965 | 2.000000 |

Table 5.4: values of $\lambda$ and $\omega$ offered to cross-validation

| $\lambda$ and $\omega$ |
|---|
| 2.000000 |
| 0.398965 |
| 0.079583 |
| 0.015875 |
| 0.003167 |
| 0.000632 |
| 0.000126 |
| 0.000025 |
| 0.000005 |
| 0.000001 |

Table 5.5: $q$ values for $q$-loss picked by cross-validation

| Data set name | Label noise (%) | | | | |
|---|---|---|---|---|---|
| | 0 | 10 | 20 | 30 | 40 |
| Long-Servedio | 0 | -0.39 | -0.24 | -0.71 | -0.55 |
| Mease-Wyner | 0 | -2.96 | -1.62 | -1.36 | 0 |
| covertype | -0.63 | -0.54 | -0.38 | -0.5 | -0.51 |
| mushrooms | 0 | -0.76 | -0.47 | -0.17 | -0.13 |
| adult9 | -0.86 | -0.53 | -0.43 | -0.53 | -0.07 |
| web8 | -0.99 | -0.46 | -0.41 | -0.19 | 0 |

Table 5.6: $C$ values for liblinear picked by cross-validation

| Data set name | Label noise (%) | | | | |
|---|---|---|---|---|---|
| | 0 | 10 | 20 | 30 | 40 |
| Long-Servedio | 0.499978 | 2.506486 | 0.499978 | 0.499978 | 0.499978 |
| Mease-Wyner | 40000.00 | 0.499978 | 315.7562 | 12.565498 | 62.99213 |
| covertype | 0.499978 | 2.506486 | 62.99213 | 1000000.0 | 12.56541 |
| mushrooms | 2.506486 | 12.56541 | 0.499978 | 0.499978 | 0.499978 |
| adult9 | 0.499978 | 62.992126 | 0.499978 | 0.499978 | 0.499978 |
| web8 | 315.7562 | 0.499978 | 12.56541 | 12.56541 | 0.499978 |

Table 5.7: $\lambda$ values picked by cross-validation for 0% label noise

| Data set name | Method | | | | |
|---|---|---|---|---|---|
| | logistic | square | smooth hinge | q-loss | sixth-order |
| Long-Servedio | 0.003167 | 0.079583 | 0.015875 | 0.015875 | 0.000001 |
| Mease-Wyner | 0.000001 | 0.000025 | 0.000001 | 0.000126 | 2.000000 |
| covertype | 0.000025 | 0.000025 | 0.000001 | 0.000025 | 0.000632 |
| mushrooms | 0.000001 | 0.000025 | 0.000632 | 0.000025 | 0.398965 |
| adult9 | 0.000001 | 0.000632 | 0.000126 | 0.003167 | 0.015875 |
| web8 | 0.000001 | 0.000005 | 0.000001 | 0.000632 | 0.015875 |

Table 5.8: $\lambda$ values picked by cross-validation for 10% label noise

| Data set name | Method | | | | |
|---|---|---|---|---|---|
| | logistic | square | smooth hinge | q-loss | sixth-order |
| Long-Servedio | 0.000005 | 2.000000 | 0.003167 | 0.015875 | 0.000001 |
| Mease-Wyner | 0.000005 | 0.000632 | 0.000005 | 0.000126 | 0.398965 |
| covertype | 0.000025 | 0.000632 | 0.000126 | 0.000001 | 0.000001 |
| mushrooms | 0.000005 | 0.000001 | 0.000005 | 0.003167 | 0.398965 |
| adult9 | 0.000632 | 0.003167 | 0.000126 | 0.015875 | 0.000632 |
| web8 | 0.000005 | 0.000126 | 0.000005 | 0.000632 | 0.015875 |

Table 5.9: $\lambda$ values picked by cross-validation for 20% label noise

| Data set name | Method | | | | |
|---|---|---|---|---|---|
| | logistic | square | smooth hinge | q-loss | sixth-order |
| Long-Servedio | 2.000000 | 2.000000 | 2.000000 | 0.000126 | 0.000001 |
| Mease-Wyner | 0.000025 | 0.000005 | 0.000126 | 0.000126 | 0.079583 |
| covertype | 0.000001 | 0.000126 | 0.000126 | 0.000001 | 0.000025 |
| mushrooms | 0.000126 | 0.000632 | 0.000025 | 0.003167 | 0.398965 |
| adult9 | 0.079583 | 0.079583 | 0.003167 | 0.015875 | 2.000000 |
| web8 | 0.000001 | 0.000001 | 0.000126 | 0.000632 | 0.015875 |

Table 5.10: $\lambda$ values picked by cross-validation for 30% label noise

| Data set name | Method | | | | |
|---|---|---|---|---|---|
| | logistic | square | smooth hinge | q-loss | sixth-order |
| Long-Servedio | 2.000000 | 2.000000 | 2.000000 | 0.003167 | 0.000001 |
| Mease-Wyner | 0.000005 | 0.000001 | 0.000005 | 0.000126 | 0.000632 |
| covertype | 0.000001 | 0.000126 | 0.000025 | 0.000025 | 0.398965 |
| mushrooms | 0.000632 | 0.003167 | 0.000632 | 0.003167 | 0.079583 |
| adult9 | 2.000000 | 0.003167 | 2.000000 | 0.003167 | 0.000632 |
| web8 | 0.000126 | 0.000001 | 0.000126 | 0.000632 | 0.000005 |

Table 5.11: $\lambda$ values picked by cross-validation for 40% label noise

| Data set name | Method | | | | |
|---|---|---|---|---|---|
| | logistic | square | smooth hinge | q-loss | sixth-order |
| Long-Servedio | 2.000000 | 2.000000 | 2.000000 | 0.003167 | 0.000632 |
| Mease-Wyner | 0.000001 | 0.000005 | 0.000025 | 0.000126 | 0.000632 |
| covertype | 0.000001 | 0.000001 | 0.000001 | 0.000001 | 0.079583 |
| mushrooms | 0.000126 | 0.000632 | 0.003167 | 0.003167 | 0.003167 |
| adult9 | 0.000126 | 0.000126 | 0.079583 | 0.000025 | 0.000025 |
| web8 | 0.015875 | 0.079583 | 0.000632 | 0.000632 | 0.000001 |

# Part II

# Adiabatic Quantum Computation

# Chapter 6

# Realizable Perturbative Gadgets for Encoding Quantum Problems

Apart from minor modifications, this chapter originally appeared as [63]:

"Hamiltonian Gadgets with Reduced Resource Requirements". Yudong Cao, Ryan Babbush, Jacob Biamonte, and Sabre Kais. *Physical Review A*. Volume 91, Number 1: 012315. 2015.

## Abstract

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. Hamiltonian gadgets, introduced around a decade ago, offer the only current means to address this requirement. Although the applications of Hamiltonian gadgets have steadily grown since their introduction, little progress has been made in overcoming the limitations of the gadgets themselves. In this 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 func-

tion 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.

## 6.1 Introduction

Although adiabatic quantum computation is known to be a universal model of quantum computation [3, 200, 216, 38] and hence, in principle equivalent to the circuit model, the mappings between an adiabatic process and an arbitrary quantum circuit require significant overhead. Currently the approaches to universal adiabatic quantum computation require implementing multiple higher order and non-commuting interactions by means of perturbative gadgets [38]. Such gadgets arose in early work on quantum complexity theory and the resources required for their implementation are the subject of this study.

Early work by Kitaev *et al.* [169] established that an otherwise arbitrary Hamiltonian restricted to have at most 5-body interactions has a ground state energy problem which is complete for the quantum analog of the complexity class NP (QMA-complete). Reducing the locality of the Hamiltonians from 5-body down to 2-body remained an open problem for a number of years. In their 2004 proof that 2-local Hamiltonian is QMA-Complete, Kempe, Kitaev and Regev formalized the idea of a perturbative gadget, which finally accomplished this task [165]. Oliveira and Terhal further reduced the problem, show-

ing completeness when otherwise arbitrary 2-body Hamiltonians were restricted to act on a square lattice [216]. The form of the simplest QMA-COMPLETE Hamiltonian is reduced to physically relevant models in [38] (see also [82]), e.g.

$$H = \sum_i h_i Z_i + \sum_{i<j} J_{ij} Z_i Z_j + \sum_{i<j} K_{ij} X_i X_j. \tag{6.1}$$

Although this model contains only physically accessible terms, programming problems into a universal adiabatic quantum computer [38] involves several types of $k$-body interactions (for bounded $k$). Reduction from $k$-body interactions to 2-body interactions is accomplished through the application of gadgets. Hamiltonian gadgets were introduced as theorem-proving tools in the context of quantum complexity theory yet their experimental realization currently offers the only currently known 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 $H$, characterized by the spectral gap $\Delta$ between its ground state subspace and excited state subspace, and a perturbation $V$ which acts on both the ancilla and the system. $V$ perturbs the ground state subspace of $H$ such that the perturbed low-lying spectrum of the gadget Hamiltonian $\widetilde{H} = H + V$ captures the spectrum of the target Hamiltonian, $H_{\text{targ}}$, up to error $\epsilon$. The purpose of a gadget depends on the form of the target Hamiltonian $H_{\text{targ}}$ and available experimental resources. 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. 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. For example the gadget which we introduce in Sec. 6.7 might fall into this category.

Apart from the physical relevance to quantum computation, gadgets have been central to many results in quantum complexity theory [54, 38, 55, 82]. Hamiltonian gadgets were also used to characterize the complexity of density functional theory [242] and are required components in current proposals related to error correction on an adiabatic quantum computer [113]. 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 [165] relied on a 2-body gadget Hamiltonian to simulate a 3-body Hamiltonian of the form $H_{\text{targ}} = H_{\text{else}} + \alpha \cdot A \otimes B \otimes C$ with three auxiliary spins in the ancilla space. Here $H_{\text{else}}$ is an arbitrary Hamiltonian that does not operate on the auxiliary spins. Further, $A$, $B$ and $C$ are unit-norm operators and $\alpha$ is the desired coupling. For such a system, it is shown that it suffices to construct $V$ with $\|V\| < \Delta/2$ to guarantee that the perturbative self-energy expansion approximates $H_{\text{targ}}$ up to error $\epsilon$ [216, 165, 54]. The gadget Hamiltonian is constructed such that in the perturbative expansion (with respect to the low energy subspace) only virtual excitations that flip all 3 ancilla bits would have non-trivial contributions in the $1^{\text{st}}$ through $3^{\text{rd}}$ order terms.

In [157] Jordan and Farhi generalized the construction in [165] to a general $k$-body to 2-body reduction. 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 [216] in their proof that 2-LOCAL HAMILTONIAN ON 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 [216]. These gadgets, which we improve in this work, find their use as the de facto standard whenever the use of gadgets is necessitated. For instance, the gadgets from [216] were used by Bravyi, DiVincenzo, Loss and Terhal [54] 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.

While recent progress in the experimental implementation of adiabatic quantum processors [43] suggests the ability to perform sophisticated adiabatic quantum computing experiments, the perturbative gadgets require very large values of $\Delta$. 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 $\Delta$. 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 $\Delta$.

### 6.1.1 Chapter structure and result summary

Previous works in the literature [165, 216, 55, 38, 54] choose $\Delta$ to be a polynomial function of $\epsilon^{-1}$ 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 $\Delta$ that guarantees error within specified $\epsilon$, which we consider here. This recipe will need to depend on three parameters: (i) the desired coupling, $\alpha$; (ii) the magnitude of the non-problematic part of the Hamiltonian, $\|H_{\text{else}}\|$; and (iii) the specified error tolerance, $\epsilon$. For simulating a target Hamiltonian up to error $\epsilon$, previous constructions

[216, 55, 54] 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 (Sec. 6.3 and 6.4) and $\Delta = \Theta(\epsilon^{-2})$ for the 3- to 2-body gadget (Sec. 6.5 and Appendix 6.8.1), showing that the physical resources required to realize the gadgets are less than previously assumed elsewhere in the literature.

In our derivation of the $\Delta$ 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 6.8.1 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 6.8.2).

The next result of this paper, described in Sec. 6.6, is a 3- to 2-body gadget construction that uses a 2-body Ising Hamiltonian with a local transverse field. This allows for the use existing flux-qubit hardware to simulate $H_{\text{targ}} = H_{\text{else}} + \alpha Z_i Z_j Z_k$ where $H_{\text{else}}$ 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 [36, 16] that can reduce a $k$-body Hamiltonian to 2-body. In this work, however, we focus on perturbative gadgets.

The final result of this paper in Sec. 6.7 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 opens the door to solving electronic structure problems on an adiabatic quantum computer.

To achieve both fast readability and completeness in presentation, each section from Sec. 6.3 to Sec. 6.7 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 paper. The **Analysis** subsections contain detailed derivations of the results in the **Summary**.

## 6.2 Perturbation theory

In our notation the spin-1/2 Pauli operators will be represented as $\{X, Y, Z\}$ with subscript indicating which spin-1/2 particle (qubit) it acts on. For example $X_2$ is a Pauli operator $X = |0\rangle\langle 1| + |1\rangle\langle 0|$ acting on the qubit represented by tensor factor 2. In the literature there are different formulations of the perturbation theory that are used when constructing and analyzing perturbative gadgets. This adds to the challenge of comparing the physical resources required among the various proposed constructions. For example, Jordan and Farhi [157] use a formulation due to Bloch, while Bravyi et al. use a formulation based on the Schrieffer-Wolff transformation [54]. Here we employ the formulation used in [165, 216]. For a review on various formulations of perturbation theory, refer to [53].

A gadget Hamiltonian $\tilde{H} = H + V$ consists of a penalty Hamiltonian $H$, which applies an energy gap onto an ancilla space, and a perturbation $V$. To explain in further detail how

the low-lying sector of the gadget Hamiltonian $\tilde{H}$ approximates the entire spectrum of a certain target Hamiltonian $H_{\text{targ}}$ with error $\epsilon$, we set up the following notations: let $\lambda_j$ and $|\psi_j\rangle$ be the $j^{\text{th}}$ eigenvalue and eigenvector of $H$ and similarly define $\tilde{\lambda}_j$ and $|\tilde{\psi}_j\rangle$ as those of $\tilde{H}$, assuming all the eigenvalues are labelled in a weakly increasing order ($\lambda_1 \leq \lambda_2 \leq \cdots$, same for $\tilde{\lambda}_j$). Using a cutoff value $\lambda_*$, let $\mathcal{L}_- = \text{span}\{|\psi_j\rangle | \forall j : \lambda_j \leq \lambda_*\}$ be the low energy subspace and $\mathcal{L}_+ = \text{span}\{|\psi_j\rangle | \forall j : \lambda_j > \lambda_*\}$ be the high energy subspace. Let $\Pi_-$ and $\Pi_+$ be the orthogonal projectors onto the subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively. For an operator $O$ we define the partitions of $O$ into the subspaces as $O_- = \Pi_- O \Pi_-$, $O_+ = \Pi_+ O \Pi_+$, $O_{-+} = \Pi_- O \Pi_+$ and $O_{+-} = \Pi_+ O \Pi_-$.

With the definitions above, one can turn to perturbation theory to approximate $\tilde{H}_-$ using $H$ and $V$. We now consider the operator-valued resolvent $\tilde{G}(z) = (zI - \tilde{H})^{-1}$. Similarly one would define $G(z) = (zI - H)^{-1}$. Note that $\tilde{G}^{-1}(z) - G^{-1}(z) = -V$ so that this allows an expansion in powers of $V$ as

$$\tilde{G} = (G^{-1} - V)^{-1} = G(I - VG)^{-1} = G + GVG + GVGVG + GVGVGVG + \cdots . \quad (6.2)$$

It is then standard to define the self-energy $\Sigma_-(z) = zI - (\tilde{G}_-(z))^{-1}$. The self-energy is important because the spectrum of $\Sigma_-(z)$ gives an approximation to the spectrum of $\tilde{H}_-$ since by definition $\tilde{H}_- = zI - \Pi_-(\tilde{G}^{-1}(z))\Pi_-$ while $\Sigma_-(z) = zI - (\Pi_- \tilde{G}(z) \Pi_-)^{-1}$. As is explained by Oliveira and Terhal [216], loosely speaking, if $\Sigma_-(z)$ is roughly constant in some range of $z$ (defined below in Theorem 6.2.1) then $\Sigma_-(z)$ is playing the role of $\tilde{H}_-$. This was formalized in [165] and improved in [216] where the following theorem is proven (as in [216] we state the case where $H$ has zero as its lowest eigenvalue and a spectral gap of $\Delta$). We use operator norm $\|\cdot\|$ which is defined as $\|M\| \equiv \max_{|\psi\rangle \in \mathcal{M}} |\langle\psi|M|\psi\rangle|$ for an operator $M$ acting on a Hilbert space $\mathcal{M}$:

**Theorem 6.2.1** (Gadget Theorem [165, 216]). *Let* $\|V\| \leq \Delta/2$ *where* $\Delta$ *is the spectral gap of* $H$ *and let the low and high spectrum of* $H$ *be separated by a cutoff* $\lambda_* = \Delta/2$. *Now let there be an effective Hamiltonian* $H_{eff}$ *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* $\Sigma_-(z)$ *has the property that* $\|\Sigma_-(z) - H_{eff}\| \leq \epsilon$, *then each eigenvalue* $\tilde{\lambda}_j$ *of* $\tilde{H}_-$ *differs to the* $j^{th}$ *eigenvalue of* $H_{eff}$, $\lambda_j$, *by at most* $\epsilon$. *In other words* $|\tilde{\lambda}_j - \lambda_j| \leq \epsilon$, $\forall j$.

To apply Theorem 6.2.1, a series expansion for $\Sigma_-(z)$ is truncated at low order for which $H_{\text{eff}}$ is approximated. The 2-body terms in $H$ and $V$ by construction can give rise to higher order terms in $H_{\text{eff}}$. For this reason it is possible to engineer $H_{\text{eff}}$ from $\Sigma_-(z)$ to approximate $H_{\text{targ}}$ up to error $\epsilon$ in the range of $z$ considered in Theorem 6.2.1 by introducing auxiliary spins and a suitable selection of 2-body $H$ and $V$. Using the series expansion of $\tilde{G}$ in Eq. 6.2, the self-energy $\Sigma_-(z) = zI - \tilde{G}_-^{-1}(z)$ can be expanded as (for further details see [165]),

$$\Sigma_-(z) = H_- + V_- + V_{-+}G_+(z)V_{+-} + V_{-+}G_+(z)V_+G_+(z)V_{+-} + \cdots . \tag{6.3}$$

The terms of $2^{\text{nd}}$ order and higher in this expansion give rise to the effective many-body interactions.

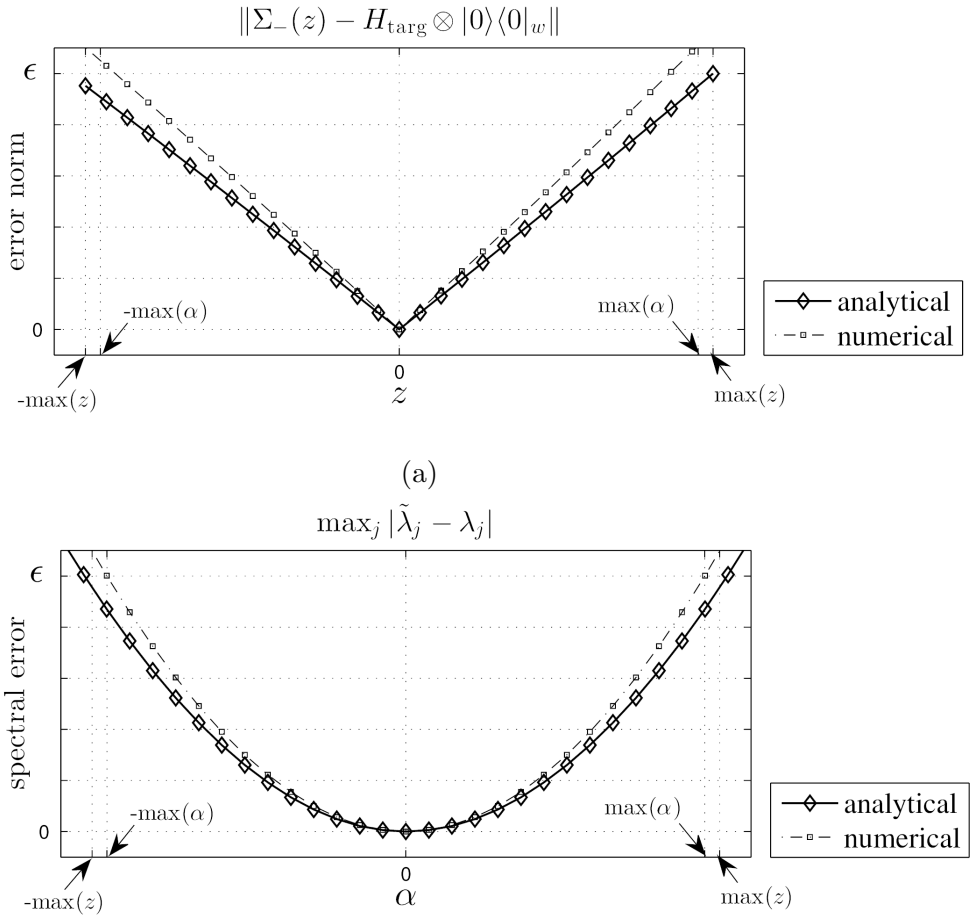



Figure 6.1: Numerical illustration of gadget theorem using a subdivision gadget. Here we use a subdivision gadget to approximate $H_{\text{targ}} = H_{\text{else}} + \alpha Z_1 Z_2$ with $\|H_{\text{else}}\| = 0$ and $\alpha \in [-1, 1]$. $\epsilon = 0.05$. "analytical" stands for the case where the value of $\Delta$ is calculated using Eq. 6.12 when $|\alpha| = 1$. "numerical" represents the case where $\Delta$ takes the value that yield the spectral error to be $\epsilon$. In (a) we let $\alpha = 1$. $z \in [-\max z, \max z]$ with $\max z = \|H_{\text{else}}\| + \max \alpha + \epsilon$. The operator $\Sigma_-(z)$ is computed up to the $3^{\text{rd}}$ order. Subplot (b) shows for every value of $\alpha$ in its range, the maximum difference between the eigenvalues $\tilde{\lambda}_j$ in the low-lying spectrum of $\tilde{H}$ and the corresponding eigenvalues $\lambda_j$ in the spectrum of $H_{\text{targ}} \otimes |0\rangle\langle 0|_w$.

## 6.3 Improved Oliveira and Terhal subdivision gadget

**Summary.** The subdivision gadget is introduced by Oliveira and Terhal [216] in their proof that 2-local Hamiltonian on square lattice is QMA-Complete. Here we show an improved lower bound for the spectral gap $\Delta$ needed on the ancilla of the gadget.

A subdivision gadget simulates a many-body target Hamiltonian $H_{\text{targ}} = H_{\text{else}} + \alpha \cdot A \otimes B$ ($H_{\text{else}}$ is a Hamiltonian of arbitrary norm, $\|A\| = 1$ and $\|B\| = 1$) by introducing an ancilla spin $w$ and applying onto it a penalty Hamiltonian $H = \Delta |1\rangle\langle 1|_w$ so that its ground state subspace $\mathcal{L}_- = \text{span}\{|0\rangle_w\}$ and its excited subspace $\mathcal{L}_+ = \text{span}\{|1\rangle_w\}$ are separated by energy gap $\Delta$. In addition to the penalty Hamiltonian $H$, we add a perturbation $V$ of the form

$$V = H_{\text{else}} + |\alpha||0\rangle\langle 0|_w + \sqrt{\frac{|\alpha|\Delta}{2}}(\text{sgn}(\alpha)A - B) \otimes X_w. \tag{6.4}$$

Hence if the target term $A \otimes B$ is $k$-local, the gadget Hamiltonian $\tilde{H} = H + V$ is at most $(\lceil k/2 \rceil + 1)$-local, accomplishing the locality reduction. Assume $H_{\text{targ}}$ acts on $n$ qubits. Prior work [216] shows that $\Delta = \Theta(\epsilon^{-2})$ is a sufficient condition for the lowest $2^n$ levels of the gadget Hamiltonian $\widetilde{H}$ to be $\epsilon$-close to the corresponding spectrum of $H_{\text{targ}}$. However, by bounding the infinite series of error terms in the perturbative expansion, we are able to obtain a tighter lower bound for $\Delta$ for error $\epsilon$. Hence we arrive at our first result (details will be presented later in this section), that it suffices to let

$$\Delta \geq \left(\frac{2|\alpha|}{\epsilon} + 1\right)(2\|H_{\text{else}}\| + |\alpha| + \epsilon). \tag{6.5}$$

In Fig. 6.2 we show numerics indicating the minimum $\Delta$ required as a function of $\alpha$ and $\epsilon$. In Fig. 6.2a the numerical results and the analytical lower bound in Eq. 6.5 show that for our subdivision gadgets, $\Delta$ can scale as favorably as $\Theta(\epsilon^{-1})$. For the subdivision gadget presented in [216], $\Delta$ scales as $\Theta(\epsilon^{-2})$. Though much less than the original assignment in

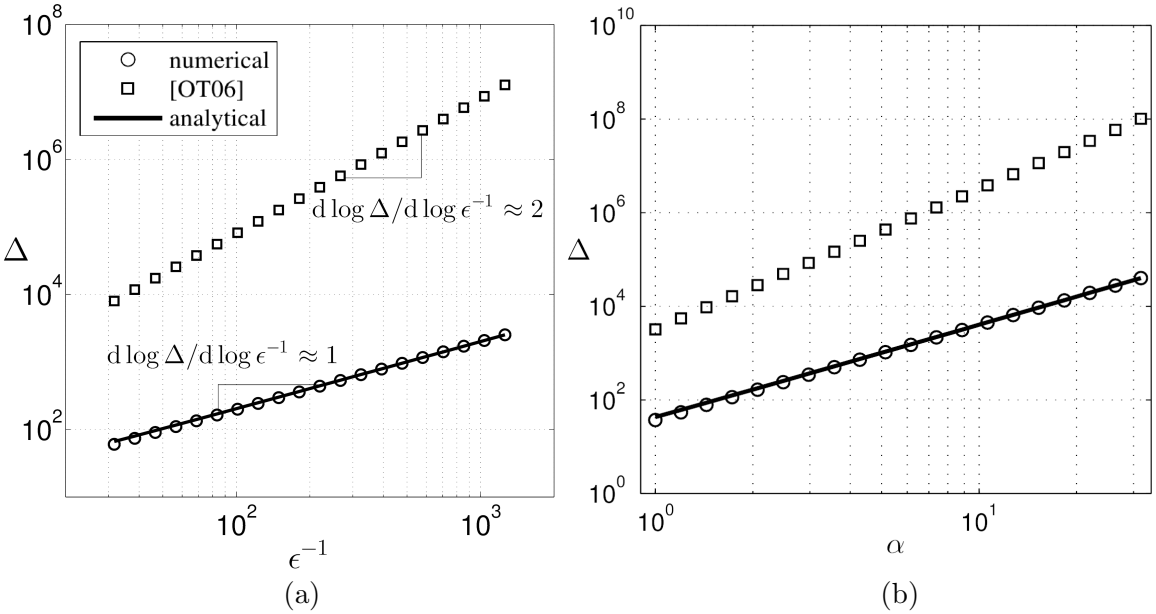



Figure 6.2: Comparison between our subdivision gadget with that of Oliveira and Terhal [216]. The data labelled as "numerical" represent the $\Delta$ values obtained from the numerical search such that the spectral error between $H_{\mathrm{targ}}$ and $\widetilde{H}_-$ is $\epsilon$. The data obtained from the calculation using Eq. 6.5 are labelled as "analytical". "[Oliveira2008]" refers to values of $\Delta$ calculated according to the assignment by Oliveira and Terhal [216]. In this example we consider $H_{\mathrm{targ}} = H_{\mathrm{else}} + \alpha Z_1 Z_2$. (a) Gap scaling with respect to $\epsilon^{-1}$. Here $\|H_{\mathrm{else}}\| = 0$ and $\alpha = 1$. (b) The gap $\Delta$ as a function of the desired coupling $\alpha$. Here $\|H_{\mathrm{else}}\| = 0$, $\epsilon = 0.05$.

[216], the lower bound of $\Delta$ in Eq. 6.5, still satisfies the condition of Theorem 6.2.1. In Fig. 6.2 we numerically find the minimum value of $\Delta$ that yields spectral error of exactly $\epsilon$.

**Analysis.** The currently known subdivision gadgets in the literature assume that the gap in the penalty Hamiltonian $\Delta$ scales as $\Theta(\epsilon^{-2})$ (see for example [216, 54]). 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 $\epsilon$. A variation of this idea will also be used to reduce the gap $\Delta$ needed in the 3- to 2-body gadget (see Sec. 6.5).

The key aspect of developing the gadget is that given $H = \Delta|1\rangle\langle 1|_w$, we need to

determine a perturbation $V$ 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 $\tilde{H} = H + V$ approximates the spectrum of the entire operator $H_{\text{targ}} \otimes |0\rangle\langle 0|_w$ up to error $\epsilon$. Here we will define $V$ and work backwards to show that it satisfies Theorem 6.2.1. We let

$$V = H_{\text{else}} + \frac{1}{\Delta}(\kappa^2 A^2 + \lambda^2 B^2) \otimes |0\rangle\langle 0|_w + (\kappa A + \lambda B) \otimes X_w \tag{6.6}$$

where $\kappa$, $\lambda$ are constants which will be determined such that the dominant contribution to the perturbative expansion which approximates $\tilde{H}_-$ gives rise to the target Hamiltonian $H_{\text{targ}} = H_{\text{else}} + \alpha \cdot A \otimes B$. In Eq. 6.6 and the remainder of the section, by slight abuse of notation, we use $\kappa A + \lambda B$ to represent $\kappa(A \otimes I_{\mathcal{B}}) + \lambda(I_{\mathcal{A}} \otimes B)$ for economy. Here $I_{\mathcal{A}}$ and $I_{\mathcal{B}}$ are identity operators acting on the subspaces $\mathcal{A}$ and $\mathcal{B}$ respectively. The partitions of $V$ in the subspaces, as defined in Sec. 6.2 are

$$\begin{gathered} V_+ = H_{\text{else}} \otimes |1\rangle\langle 1|_w, \quad V_- = \left(H_{\text{else}} + \frac{1}{\Delta}(\kappa^2 A^2 + \lambda^2 B^2) I\right) \otimes |0\rangle\langle 0|_w, \\ V_{-+} = (\kappa A + \lambda B) \otimes |0\rangle\langle 1|_w, \quad V_{+-} = (\kappa A + \lambda B) \otimes |1\rangle\langle 0|_w. \end{gathered} \tag{6.7}$$

We would like to approximate the target Hamiltonian $H_{\text{targ}}$ and so expand the self-energy in Eq. 6.3 up to 2$^{\text{nd}}$ order. Note that $H_- = 0$ and $G_+(z) = (z - \Delta)^{-1}|1\rangle\langle 1|_w$. Therefore the self energy $\Sigma_-(z)$ can be expanded as

$$\Sigma_-(z) = V_- + \frac{1}{z - \Delta} V_{-+} V_{+-} + \sum_{k=1}^{\infty} \frac{V_{-+} V_+^k V_{+-}}{(z - \Delta)^{k+1}} \tag{6.8}$$

$$= \underbrace{\left(H_{\text{else}} - \frac{2\kappa\lambda}{\Delta} A \otimes B\right) \otimes |0\rangle\langle 0|_w}_{H_{\text{eff}}} + \underbrace{\frac{z}{\Delta(z - \Delta)}(\kappa A + \lambda B)^2 \otimes |0\rangle\langle 0|_w + \sum_{k=1}^{\infty} \frac{V_{-+} V_+^k V_{+-}}{(z - \Delta)^{k+1}}}_{\text{error term}}.$$

By selecting $\kappa = \text{sgn}(\alpha)(|\alpha|\Delta/2)^{1/2}$ and $\lambda = -(|\alpha|\Delta/2)^{1/2}$, the leading order term in $\Sigma_-(z)$ becomes $H_{\text{eff}} = H_{\text{targ}} \otimes |0\rangle\langle 0|_w$. We must now show that the condition of Theorem 6.2.1 is

satisfied i.e. for a small real number $\epsilon > 0$, $\|\Sigma_-(z) - H_{\text{eff}}\| \leq \epsilon, \forall z \in [\min z, \max z]$ where $\max z = \|H_{\text{else}}\| + |\alpha| + \epsilon = -\min z$. Essentially this amounts to choosing a value of $\Delta$ to cause the error term in Eq. 6.8 to be $\leq \epsilon$. In order to derive a tighter lower bound for $\Delta$, we bound the norm of the error term in Eq. 6.8 by letting $z \mapsto \max z$ and from the triangle inequality for operator norms:

$$\left\| \frac{z}{\Delta(z-\Delta)} (\kappa A + \lambda B)^2 \otimes |0\rangle\langle 0|_w \right\| \leq \frac{\max z}{\Delta(\Delta - \max z)} \cdot 4\kappa^2 = \frac{2|\alpha| \max z}{\Delta - \max z} \tag{6.9}$$

$$\left\| \sum_{k=1}^{\infty} \frac{V_{-+} V_+^k V_{+-}}{(z-\Delta)^{k+1}} \right\| \leq \sum_{k=1}^{\infty} \frac{\|V_{-+}\| \cdot \|V_+\|^k \cdot \|V_{+-}\|}{(\Delta - \max z)^{k+1}}$$

$$\leq \sum_{k=1}^{\infty} \frac{2|\kappa| \cdot \|H_{\text{else}}\|^k \cdot 2|\kappa|}{(\Delta - \max z)^{k+1}} = \sum_{k=1}^{\infty} \frac{2|\alpha|\Delta \|H_{\text{else}}\|^k}{(\Delta - \max z)^{k+1}}.$$

Using $H_{\text{eff}} = H_{\text{targ}} \otimes |0\rangle\langle 0|_w$, from (6.8) we see that

$$\|\Sigma_-(z) - H_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \frac{2|\alpha| \max z}{\Delta - \max z} + \sum_{k=1}^{\infty} \frac{2|\alpha|\Delta \|H_{\text{else}}\|^k}{(\Delta - \max z)^{k+1}} \tag{6.10}$$

$$= \frac{2|\alpha| \max z}{\Delta - \max z} + \frac{2|\alpha|\Delta}{\Delta - \max z} \cdot \frac{\|H_{\text{else}}\|}{\Delta - \max z - \|H_{\text{else}}\|}. \tag{6.11}$$

Here going from Eq. 6.10 to Eq. 6.11 we have assumed the convergence of the infinite series in Eq. 6.10, which adds the reasonable constraint that $\Delta > |\alpha| + \epsilon + 2\|H_{\text{else}}\|$. To ensure that $\|\Sigma_-(z) - H_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \epsilon$ it is sufficient to let expression Eq. 6.11 be $\leq \epsilon$, which implies that

$$\Delta \geq \left( \frac{2|\alpha|}{\epsilon} + 1 \right) (|\alpha| + \epsilon + 2\|H_{\text{else}}\|) \tag{6.12}$$

which is $\Theta(\epsilon^{-1})$, a tighter bound than $\Theta(\epsilon^{-2})$ in the literature [54, 165, 216]. This bound is illustrated with a numerical example (Fig. 6.1). From the data labelled as "analytical" in Fig. 6.1a we see that the error norm $\|\Sigma_-(z) - H_{\text{eff}}\|$ is within $\epsilon$ for all $z$ considered in the range, which satisfies the condition of the theorem for the chosen example. In Fig. 6.1b, the data labelled "analytical" show that the spectral difference between $\tilde{H}_-$ and

$H_{\text{eff}} = H_{\text{targ}} \otimes |0\rangle\langle 0|_w$ is indeed within $\epsilon$ as the theorem promises. Furthermore, note that the condition of Theorem 6.2.1 is only sufficient, which justifies why in Fig. 6.1b for $\alpha$ values at $\max \alpha$ and $\min \alpha$ the spectral error is strictly below $\epsilon$. This indicates that an even smaller $\Delta$, although below the bound we found in Eq. 6.12 to satisfy the theorem, could still yield the spectral error within $\epsilon$ for all $\alpha$ values in the range. The smallest value $\Delta$ can take would be one such that the spectral error is exactly $\epsilon$ when $\alpha$ is at its extrema. We numerically find this $\Delta$ (up to numerical error which is less than $10^{-5}\epsilon$) and as demonstrated in Fig. 6.1b, the data labelled "numerical" shows that the spectral error is indeed $\epsilon$ at $\max(\alpha)$ and $\min(\alpha)$, yet in Fig. 6.1a the data labelled "numerical" shows that for some $z$ in the range the condition of the Theorem 6.2.1, $\|\Sigma_-(z) - H_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \epsilon$, no longer holds. In Fig. 6.1 we assume that $\epsilon$ is kept constant. In Fig. 6.2a we compute both analytical and numerical $\Delta$ values for different values of $\epsilon$.

*Comparison with Oliveira and Terhal [216].* We also compare our $\Delta$ assignment with the subdivision gadget by Oliveira and Terhal [216], where given a target Hamiltonian $H_{\text{targ}} = H_{\text{else}} + Q \otimes R$ it is assumed that $Q$ and $R$ are operators with finite norm operating on two separate spaces $\mathcal{A}$ and $\mathcal{B}$.

The construction of the subdivision gadget in [216] is the same as the construction presented earlier: introduce an ancillary qubit $w$ with energy gap $\Delta$, then the unperturbed Hamiltonian is $H = \Delta|1\rangle\langle 1|_w$. In [216] they add a perturbation $V$ that takes the form of [216, Eq. 15]

$$V = H'_{\text{else}} + \sqrt{\frac{\Delta}{2}}(-Q + R) \otimes X_w \tag{6.13}$$

where $H'_{\text{else}} = H_{\text{else}} + Q^2/2 + R^2/2$. Comparing the form of Eq. 6.13 and Eq. 6.6 we can see that if we redefine $Q = \sqrt{|\alpha|}A$ and $R = \sqrt{|\alpha|}B$, the gadget formulation is identical to

our subdivision gadget approximating $H_{\text{targ}} = H_{\text{else}} + \alpha A \otimes B$ with $\alpha > 0$. In the original work $\Delta$ is chosen as [216, Eq. 20]

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

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

$$\Delta \geq \frac{(\|H_{\text{else}} + |\alpha| I\| + \sqrt{2|\alpha|})^6}{\epsilon^2}. \tag{6.14}$$

In Fig. 6.2a we compare the lower bound in Eq. 6.14 with our lower bound in Eq. 6.12 and the numerically optimized $\Delta$ described earlier.

## 6.4 Parallel subdivision and $k$- to 3-body reduction

**Summary.** Applying subdivision gadgets iteratively one can reduce a $k$-body Hamiltonian $H_{\text{targ}} = H_{\text{else}} + \alpha \bigotimes_{i=1}^{k} \sigma_i$ to 3-body. Here each $\sigma_i$ is a single spin Pauli operator. Initially, the term $\bigotimes_{i=1}^{k} \sigma_i$ can be broken down into $A \otimes B$ where $A = \bigotimes_{i=1}^{r} \sigma_i$ and $B = \bigotimes_{i=r+1}^{k} \sigma_i$. 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 $i^{\text{th}}$ 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 $\widetilde{H}_i$ and $\widetilde{H}_{i-1}$ is within $\epsilon$. From the recurrence relation $\Delta_i = \Theta(\epsilon^{-1} \Delta_{i-1}^{3/2})$, we hence were able to show a quadratic improvement over previous $k$-body constructions [54].

**Analysis.** The concept of parallel application of gadgets has been introduced in [216, 165]. The idea of using subdivision gadgets for iteratively reducing a $k$-body Hamiltonian to 3-body has been mentioned in [216, 54]. 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 $\epsilon$. For the purpose of presentation, let us define the notions of "parallel" and "series" gadgets in the following remarks.

**Remark 6.4.1** (Parallel gadgets)**.** *Parallel application of gadgets refers to using gadgets on multiple terms $H_{targ,i}$ in the target Hamiltonian $H_{targ} = H_{else} + \sum_{i=1}^{m} H_{targ,i}$ concurrently. Here one will introduce $m$ ancilla spins $w_1, \cdots, w_m$ and the parallel gadget Hamiltonian takes the form of $\tilde{H} = \sum_{i=1}^{m} H_i + V$ where $H_i = \Delta|1\rangle\langle 1|_{w_i}$ and $V = H_{else} + \sum_{i=1}^{m} V_i$. $V_i$ is the perturbation term of the gadget applied to $H_{targ,i}$.*

**Remark 6.4.2** (Serial gadgets)**.** *Serial application of gadgets refers to using gadgets sequentially. Suppose the target Hamiltonian $H_{targ}$ is approximated by a gadget Hamiltonian $\tilde{H}^{(1)}$ such that $\tilde{H}^{(1)}_{-}$ approximates the spectrum of $H_{targ}$ up to error $\epsilon$. If one further applies onto $\tilde{H}^{(1)}$ another gadget and obtains a new Hamiltonian $\tilde{H}^{(2)}$ whose low-lying spectrum captures the spectrum of $\tilde{H}^{(1)}$, we say that the two gadgets are applied in series to reduce $H_{targ}$ to $\tilde{H}^{(2)}$.*

Based on Remark 6.4.1, a parallel subdivision gadget deals with the case where $H_{\mathrm{targ,i}} = \alpha_i A_i \otimes B_i$. $\alpha_i$ is a constant and $A_i$, $B_i$ are unit norm Hermitian operators that act on separate spaces $\mathcal{A}_i$ and $\mathcal{B}_i$. Note that with $H_i = \Delta|1\rangle\langle 1|_{w_i}$ for every $i \in \{1, 2, \cdots, m\}$ we have the total penalty Hamiltonian $H = \sum_{i=1}^{m} H_i = \sum_{x\in\{0,1\}^m} h(x)\Delta|x\rangle\langle x|$ 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 $\Delta$. The operator-valued resolvent $G$ for the penalty Hamiltonian is (by definition in Sec. 6.2)

$$G(z) = \sum_{x\in\{0,1\}^m} \frac{1}{z - h(x)\Delta} |x\rangle\langle x|. \tag{6.15}$$

The perturbation Hamiltonian $V$ is defined as

$$V = H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2 A_i^2 + \lambda_i^2 B_i^2) + \sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i) \otimes X_{u_i} \tag{6.16}$$

where the coefficients $\kappa_i$ and $\lambda_i$ are defined as $\kappa_i = \text{sgn}(\alpha_i)\sqrt{|\alpha_i|\Delta/2}, \lambda_i = -\sqrt{|\alpha_i|\Delta/2}$.

Define $P_- = |0\rangle^{\otimes m}\langle 0|^{\otimes m}$ and $P_+ = I - P_-$. Then if $H_{\text{targ}}$ acts on the Hilbert space $\mathcal{M}$, $\Pi_- = I_{\mathcal{M}} \otimes P_-$ and $\Pi_+ = I_{\mathcal{M}} \otimes P_+$. Comparing Eq. 6.16 with Eq. 6.6 we see that the projector to the low-lying subspace $|0\rangle\langle 0|_w$ in Eq. 6.6 is replaced by an identity $I$ in Eq. 6.16. This is because in the case of $m$ parallel gadgets $P_-$ cannot be realized with only 2-body terms when $m \geq 3$.

The partition of $V$ in the subspaces are

$$\begin{aligned} V_- &= \left(H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2 A_i^2 + \lambda_i^2 B_i^2)\right) \otimes P_-, \quad V_+ = \left(H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2 A_i^2 + \lambda_i^2 B_i^2)\right) \otimes P_+ \\ V_{-+} &= \sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i) \otimes P_- X_{u_i} P_+, \qquad V_{+-} = \sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i) \otimes P_+ X_{u_i} P_-. \end{aligned} \tag{6.17}$$

The self-energy expansion in Eq. 6.3 then becomes

$$\begin{aligned} \Sigma_-(z) &= \left(H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2 A_i^2 + \lambda_i^2 B_i^2)\right) \otimes P_- + \frac{1}{z-\Delta}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)^2 \otimes P_- \\ &+ \sum_{k=1}^{\infty} V_{-+}(G_+ V_+)^k G_+ V_{+-}. \end{aligned} \tag{6.18}$$

Rearranging the terms we have

$$\Sigma_-(z) = \underbrace{\left(H_{\text{else}} + \sum_{i=1}^{m}\left(-\frac{2\kappa_i\lambda_i}{\Delta}A_i \otimes B_i\right)\right)\otimes P_-}_{H_{\text{eff}}} + \underbrace{\left(\frac{1}{\Delta}+\frac{1}{z-\Delta}\right)\sum_{i=1}^{m}(\kappa_i^2 A_i^2 + \lambda_i^2 B_i^2)\otimes P_-}_{E_1}$$

$$+ \underbrace{\left(\frac{1}{\Delta}+\frac{1}{z-\Delta}\right)\sum_{i=1}^{m} 2\kappa_i\lambda_i A_i \otimes B_i \otimes P_-}_{E_2} + \underbrace{\sum_{k=1}^{\infty} V_{-+}(G_+V_+)^k G_+ V_{+-}}_{E_3}[0.1in] \tag{6.19}$$

where the term $H_{\text{eff}} = H_{\text{targ}} \otimes P_-$ is the effective Hamiltonian that we would like to obtain from the perturbative expansion and $E_1$, $E_2$, and $E_3$ are error terms. Theorem 6.2.1 states that for $z \in [-\max(z), \max(z)]$, if $\|\Sigma_-(z) - H_{\text{targ}} \otimes P_-\| \leq \epsilon$ then $\tilde{H}_-$ approximates the spectrum of $H_{\text{targ}} \otimes P_-$ by error at most $\epsilon$. Similar to the triangle inequality derivation shown in (6.9), to derive a lower bound for $\Delta$, let $z \mapsto \max(z) = \|H_{\text{else}}\| + \sum_{i=1}^{m}|\alpha_i| + \epsilon$ and the upper bounds of the error terms $E_1$ and $E_2$ can be found as

$$\begin{aligned} \|E_1\| &\leq \frac{\max(z)}{\Delta - \max(z)}\sum_{i=1}^{m}|\alpha_i| \leq \frac{\max(z)}{\Delta-\max(z)}\left(\sum_{i=1}^{m}|\alpha_i|^{1/2}\right)^2 \\ \|E_2\| &\leq \frac{\max(z)}{\Delta - \max(z)}\left(\sum_{i=1}^{m}|\alpha_i|^{1/2}\right)^2. \end{aligned} \tag{6.20}$$

From the definition in Eq. 6.15 we see that $\|G_+(z)\| \leq \frac{1}{\Delta - \max(z)}$. Hence the norm of $E_3$ can be bounded by

$$\begin{aligned} \|E_3\| &\leq \sum_{k=1}^{\infty}\frac{\|\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)\|^2 \|H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i^2 A_i^2 + \lambda_i^2 B_i^2) I\|^k}{(\Delta - \max(z))^{k+1}} \\ &\leq \sum_{k=1}^{\infty}\frac{2\Delta(\sum_{i=1}^{m}|\alpha_i|^{1/2})^2(\|H_{\text{else}}\| + \sum_{i=1}^{m}|\alpha_i|)^k}{(\Delta-\max(z))^{k+1}} \\ &= \frac{2\Delta(\sum_{i=1}^{m}|\alpha_i|^{1/2})^2}{\Delta - \max(z)}\frac{\|H_{\text{else}}\| + \sum_{i=1}^{m}|\alpha_i|}{\Delta - \max(z) - (\|H_{\text{else}}\| + \sum_{i=1}^{m}|\alpha_i|)}. \end{aligned} \tag{6.21}$$

Similar to the discussion in Sec. 6.3, to ensure that $\|\Sigma_-(z) - H_{\text{targ}} \otimes P_-\| \leq \epsilon$, which is the

condition of Theorem 6.2.1, it is sufficient to let $\|E_1\| + \|E_2\| + \|E_3\| \leq \epsilon$:

$$\begin{aligned} \|E_1\| + \|E_2\| + \|E_3\| &\leq \frac{2\max(z)}{\Delta - \max(z)} \left( \sum_{i=1}^{m} |\alpha_i|^{1/2} \right)^2 \\ &+ \frac{2\Delta(\sum_{i=1}^{m} |\alpha_i|^{1/2})^2}{\Delta - \max(z)} \cdot \frac{\|H_{\text{else}}\| + \sum_{i=1}^{m} |\alpha_i|}{\Delta - \max(z) - (\|H_{\text{else}}\| + \sum_{i=1}^{m} |\alpha_i|)} \\ &= \frac{2(\sum_{i=1}^{m} |\alpha_i|^{1/2})^2 (\max(z) + \|H_{\text{else}}\| + \sum_{i=1}^{m} |\alpha_i|)}{\Delta - \max(z) - (\|H_{\text{else}}\| + \sum_{i=1}^{m} |\alpha_i|)} \leq \epsilon \end{aligned} \tag{6.22}$$

where we find the lower bound of $\Delta$ for parallel subdivision gadget

$$\Delta \geq \left[ \frac{2(\sum_{i=1}^{m} |\alpha_i|^{1/2})^2}{\epsilon} + 1 \right] (2\|H_{\text{else}}\| + 2\sum_{i=1}^{m} |\alpha_i| + \epsilon). \tag{6.23}$$

Note that if one substitutes $m = 1$ into Eq. 6.23 the resulting expression is a lower bound that is less tight than that in Eq. 6.12. This is because of the difference in the perturbation $V$ between Eq. 6.16 and Eq. 6.6 which is explained in the text preceding Eq. 6.17. Also we observe that the scaling of this lower bound for $\Delta$ 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 [216, 165, 54] that subdivision gadgets can be applied to multiple terms in parallel and the scaling of the gap $\Delta$ 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 following iterative scheme 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 $i^{\text{th}}$ iteration and subscript $_i$ for labelling objects

within the same iteration):

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

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

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

Step 2: Partition the rest of the terms in $\tilde{H}^{(i-1)}$ into $\alpha_1^{(i)} A_1^{(i)} \otimes B_1^{(i)}$, $\alpha_2^{(i)} A_2^{(i)} \otimes B_2^{(i)}$, $\cdots$, $\alpha_m^{(i)} A_m^{(i)} \otimes B_m^{(i)}$. Here $\alpha_j^{(i)}$ are coefficients.

Step 3: Introduce $m$ ancilla qubits $w_1^{(i)}$, $w_2^{(i)}$, $\cdots w_m^{(i)}$ and construct $\tilde{H}^{(i)}$ using the parallel subdivision gadget. Let $P_-^{(i)} = |0\cdots 0\rangle\langle 0\cdots 0|_{w_1^{(i)}\cdots w_m^{(i)}}$.

Define $\Pi_-^{(i)} = I_{\mathcal{M}^{(i)}} \otimes P_-^{(i)}$.

3.1: Apply the penalty Hamiltonian $H^{(i)} = \sum_{x\in\{0,1\}^m} h(x)\Delta^{(i)}|x\rangle\langle x|$.

Here $\Delta^{(i)}$ is calculated by the lower bound in Eq. 6.23.

3.2: Apply the perturbation $V^{(i)} =$

$H_{\text{else}}^{(i)} + \sum_{j=1}^m \sqrt{\frac{|\alpha_j^{(i)}|\Delta^{(i)}}{2}}(\text{sgn}(\alpha_j^{(i)})A_j^{(i)} - B_j^{(i)}) \otimes X_{w_j^{(i)}} + \sum_{j=1}^m |\alpha_j^{(i)}| I$.

3.3: $\tilde{H}^{(i)} = H^{(i)} + V^{(i)}$ acts on the space $\mathcal{M}^{(i)}$ and the maximum spectral difference between $\tilde{H}_-^{(i)} = \Pi_-^{(i)}\tilde{H}^{(i)}\Pi_-^{(i)}$ and $\tilde{H}^{(i-1)} \otimes P_-^{(i)}$ is at most $\epsilon$.

$i \to i+1$

**end**

(6.24)

We could show that after $s$ iterations, the maximum spectral error between $\Pi_-^{(s)}\tilde{H}^{(s)}\Pi_-^{(s)}$ and $\tilde{H}^{(0)}\bigotimes_{i=1}^s P_-^{(s)}$ is guaranteed to be within $s\epsilon$. Suppose we would like to make target Hamiltonian $\tilde{H}_0$, we construct a gadget $\tilde{H} = H^{(1)} + V^{(1)}$ according to algorithm (6.24), such that $|\lambda(\tilde{H}^{(1)}) - \lambda(\tilde{H}^{(0)})| \leq \epsilon$ for low-lying eigenvalues $\lambda(\cdot)$. Note that in a precise sense we should write $|\lambda(\Pi_-^{(1)}\tilde{H}^{(1)}\Pi_-^{(1)}) - \lambda(\tilde{H}^{(0)} \otimes P_-^{(0)})|$. Since the projectors $\Pi_-^{(i)}$ and $P_-^{(i)}$ do

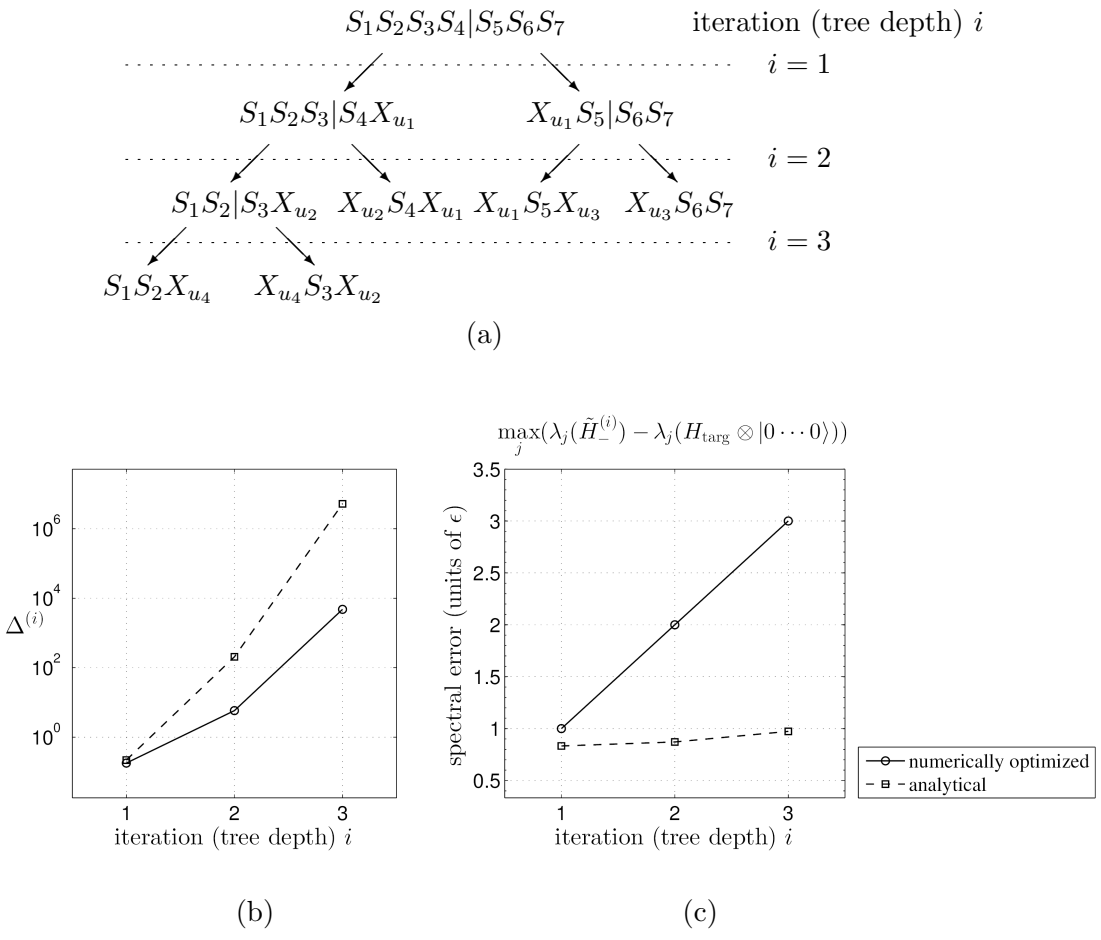



Figure 6.3: (a) Reduction tree diagram for reducing a 7-body term to 3-body using parallel subdivision gadgets. Each $S_i$ is a single-qubit Pauli operator acting on qubit $i$. The vertical lines | show where the subdivisions are made at each iteration to each term. (b) An example where we consider the target Hamiltonian $H_{\text{targ}} = \alpha S_1S_2S_3S_4S_5S_6S_7$ with $\alpha = 5 \times 10^{-3}$, $S_i = X_i$, $\forall i \in \{1, 2, \cdots, 7\}$, and reduce it to 3-body according to (a) up to error $\epsilon = 5\times 10^{-4}$. 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 $\tilde{H}^{(i)}$ and the target Hamiltonian $H_{\text{targ}}$. 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 $\Pi_-^{(i)}\tilde{H}^{(i)}\Pi_-^{(i)}$ and $\tilde{H}^{(i-1)} \otimes P_-^{(i)}$ is $\epsilon$. For definitions of $\Delta^{(i)}$, $\tilde{H}^{(i)}$, $\Pi_-^{(i)}$ and $P_-^{(i)}$, see Eq. 6.24. Those labelled as 'analytical' correspond to cases where each iteration uses the gap bound derived in Eq. 6.23.

not affect the low-lying spectrum of $\tilde{H}^{(i)}$ and $\tilde{H}^{(i-1)}$, for simplicity and clarity we write only $\tilde{H}^{(i-1)}$ and $\tilde{H}^{(i)}$. After $\tilde{H}^{(1)}$ is introduced, according to algorithm (6.24) the second gadget $\tilde{H}^{(2)}$ is constructed by considering the *entire* $\tilde{H}^{(1)}$ as the new target Hamiltonian and introducing ancilla particles with unperturbed Hamiltonian $H^{(2)}$ and perturbation $V^{(2)}$ such that the low-energy spectrum of $\tilde{H}^{(2)}$ approximates the spectrum of $\tilde{H}^{(1)}$ up to error $\epsilon$. That is, $|\lambda(\tilde{H}^{(1)}) - \lambda(\tilde{H}^{(2)})| \leq \epsilon$. With the serial application of gadgets we produce a sequence of Hamiltonians $\tilde{H}^{(0)} \to \tilde{H}^{(1)} \to \tilde{H}^{(2)} \to \cdots \to \tilde{H}^{(k)}$ where $\tilde{H}^{(0)}$ is the target Hamiltonian and each subsequent gadget Hamiltonian $\tilde{H}^{(i)}$ captures the *entire* previous gadget $\tilde{H}^{(i-1)}$ in its low-energy sector with $|\lambda(\tilde{H}^{(i)}) - \lambda(\tilde{H}^{(i-1)})| \leq \epsilon$. Hence to bound the spectral error between the last gadget $\tilde{H}^{(k)}$ and the target Hamiltonian $\tilde{H}^{(0)}$ we could use triangle inequality: $|\lambda(\tilde{H}^{(s)}) - \lambda(\tilde{H}^{(0)})| \leq |\lambda(\tilde{H}^{(s)}) - \lambda(\tilde{H}^{(s-1)})| + \cdots + |\lambda(\tilde{H}^{(1)}) - \lambda(\tilde{H}^{(0)})| \leq s\epsilon$.

*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

$$f(k) = \begin{cases} f\left(\frac{k}{2}+1\right)+1 & k \text{ even} \\ f\left(\frac{k+1}{2}+1\right)+1 & k \text{ odd} \end{cases} \tag{6.25}$$

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 $\tilde{H}^{(s)}$ and $\tilde{H}^{(0)}$ 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,

$$\Delta^{(s)} = \Theta\left(\epsilon^{-2[(3/2)^{s-1}-1]}(\Delta^{(1)})^{(3/2)^{s-1}}\right). \tag{6.26}$$

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

$$\Delta^{(s)} = \Theta\left(\epsilon^{-3(\frac{1}{2}\lceil k-2 \rceil)^{\log_2(3/2)}-2}\right) = \Theta\left(\epsilon^{-\text{poly}(k)}\right) \tag{6.27}$$

accumulating exponentially as a function of $k$. The exponential nature of the scaling with respect to $k$ agrees with results by Bravyi et al. [54]. However, in our construction, 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 [54], which is $\Delta^{(i+1)} = \Theta(\epsilon^{-2}(\Delta^{(i)})^3)$.

*Qubit cost.* Based on the reduction scheme described in Eq. 6.24 (illustrated in Fig. 6.3a 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 $S_1 S_2 \cdots S_k$ (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 Eq. 6.24. At each iteration, if we describe every individual subdivision gadget by a vertical line $|$ at the location where the partition is made, for example $S_1S_2S_3S_4|S_5S_6S_7$ in the case of the first iteration in Fig. 6.3a, then after $\lceil \log_2(k-2) \rceil$ iterations all the partitions made to the $k$-body term can be described as $S_1S_2|S_3|S_4|\cdots|S_{k-2}|S_{k-1}S_k$. 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 Eq. 6.24 on a target Hamiltonian $H_{\text{targ}} = \alpha S_1S_2S_3S_4S_5S_6S_7$. Here we let $S_i = X_i$, $\forall i \in \{1, 2, \cdots, 7\}$, $\epsilon = 5 \times 10^{-4}$ and $\alpha = 5 \times 10^{-3}$. During each iteration the values of $\Delta^{(i)}$ are

assigned according to the lower bound in Eq. 6.23. From Fig. 6.3c we can see that the lower bounds are sufficient for keeping the total spectral error between $\tilde{H}_-^{(3)}$ and $\tilde{H}^{(0)} \bigotimes_{i=1}^{3} P_-^{(i)}$ within $3\epsilon$. Furthermore, numerical search is also used at each iteration to find the minimum value of $\Delta^{(i)}$ so that the spectral error between $\Pi_-^{(i)} \tilde{H}^{(i)} \Pi_-^{(i)}$ and $\tilde{H}^{(i-1)} \bigotimes_{j=1}^{i} P_-^{(j)}$ is $\epsilon$. The numerically found gaps $\Delta^{(i)}$ are much smaller than their analytical counterparts at each iteration (Fig. 6.3b), at the price that the error is larger (Fig. 6.3c). In both the numerical and the analytical cases, the error appears to accumulate linearly as the iteration proceeds.

## 6.5 Improved Oliveira and Terhal 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 [165, 216, 54]. Consider 3-body target Hamiltonian of the form $H_{\text{targ}} = H_{\text{else}} + \alpha A \otimes B \otimes C$. Here $A$, $B$ and $C$ 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 [216] and also used in Bravyi, DiVincenzo, Loss and Terhal [54]. To accomplish the 3- to 2-body reduction, we introduce an ancilla spin $w$ and apply a penalty Hamiltonian $H = \Delta|1\rangle\langle 1|_w$. We then add a perturbation $V$ of form,

$$V = H_{\text{else}} + \mu C \otimes |1\rangle\langle 1|_w + (\kappa A + \lambda B) \otimes X_w + V_1 + V_2 \tag{6.28}$$

where $V_1$ and $V_2$ are 2-local compensation terms (details presented later in this section):

$$\begin{aligned} V_1 &= \frac{1}{\Delta}(\kappa^2 + \lambda^2)|0\rangle\langle 0|_w + \frac{2\kappa\lambda}{\Delta} A \otimes B - \frac{1}{\Delta^2}(\kappa^2 + \lambda^2)\mu C \otimes |0\rangle\langle 0|_w \\ V_2 &= -\frac{2\kappa\lambda}{\Delta^3}\text{sgn}(\alpha)\Big[(\kappa^2 + \lambda^2)|0\rangle\langle 0|_w + 2\kappa\lambda A \otimes B\Big]. \end{aligned} \tag{6.29}$$

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 $\Delta$, the low-lying spectrum of the gadget Hamiltonian $\widetilde{H}$ captures

the entire spectrum of $H_{\text{targ}}$ up to arbitrary error $\epsilon$. In the construction of [54] it is shown that $\Delta = \Theta(\epsilon^{-3})$ is sufficient. In [165], $\Delta = \Theta(\epsilon^{-3})$ is also assumed, though the construction of $V$ is slightly different from Eq. 6.28. By adding terms in $V$ 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:

$$\Delta \geq \frac{1}{4}(-b + \sqrt{b^2 - 4c})^2 \tag{6.30}$$

where $b$ and $c$ are defined as

$$\begin{aligned} b &= -\left[\xi + \frac{2^{4/3}\alpha^{2/3}}{\epsilon}(\max z + \eta + \xi^2)\right] \\ c &= -\left(1 + \frac{2^{4/3}\alpha^{2/3}}{\epsilon}\xi\right)(\max z + \eta) \end{aligned} \tag{6.31}$$

with $\max z = \|H_{\text{else}}\| + |\alpha| + \epsilon$, $\eta = \|H_{\text{else}}\| + 2^{2/3}\alpha^{4/3}$ and $\xi = 2^{-1/3}\alpha^{1/3} + 2^{1/3}\alpha^{2/3}$. From Eq. 6.30 we can see the lower bound to $\Delta$ is $\Theta(\epsilon^{-2})$. Our improvement results in a power of $\epsilon^{-1}$ reduction in the gap. For the dependence of $\Delta$ on $\|H_{\text{else}}\|$, $\alpha$ and $\epsilon^{-1}$ for both the original [216] and the optimized case, see Fig. 6.4. Results show that the bound in Eq. 6.30 is tight with respect to the minimum $\Delta$ numerically found that yields the spectral error between $\tilde{H}_-$ and $H_{\text{targ}} \otimes |0\rangle\langle 0|_w$ to be $\epsilon$.

**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 [216, 54] and argue that $\Delta = \Theta(\epsilon^{-3})$ is required for yielding a spectral error between $\tilde{H}$ and $H_{\text{eff}}$ within $\epsilon$ using this construction.

In the improved construction we define the perturbation $V$ as in Eq. 6.28. 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 $\Delta$ in the coefficients are optimal, we introduce a parameter $r$

such that

$$\kappa = \mathrm{sgn}(\alpha)\left(\frac{\alpha}{2}\right)^{1/3}\Delta^r, \qquad \lambda = \left(\frac{\alpha}{2}\right)^{1/3}\Delta^r, \qquad \mu = \left(\frac{\alpha}{2}\right)^{1/3}\Delta^{2-2r}. \tag{6.32}$$

It is required that $\|V\| \leq \Delta/2$ (Theorem 6.2.1) 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, $V_-$, $V_+$, $V_{-+}$, $V_{+-}$ can be calculated as the following:

$$\begin{aligned}
V_- &= \left[H_{\text{else}} + \frac{1}{\Delta}(\kappa A + \lambda B)^2 - \frac{1}{\Delta}(\kappa^2+\lambda^2)\mu C - \frac{2\kappa\lambda}{\Delta^3}\mathrm{sgn}(\alpha)(\kappa A + \lambda B)^2\right] \otimes |0\rangle\langle 0|_w \\
V_+ &= \left[H_{\text{else}} + \mu C + \frac{2\kappa\lambda}{\Delta} A\otimes B - \frac{4\kappa^2\lambda^2}{\Delta^3}\mathrm{sgn}(\alpha) A\otimes B\right] \otimes |1\rangle\langle 1|_w \\
V_{-+} &= (\kappa A + \lambda B)\otimes |0\rangle\langle 1|_w \\
V_{+-} &= (\kappa A + \lambda B)\otimes |1\rangle\langle 0|_w.
\end{aligned} \tag{6.33}$$

The self-energy expansion, referring to Eq. 6.3, becomes

$$\Sigma_-(z) = V_- + \frac{1}{z-\Delta}V_{-+}V_{+-} + \frac{1}{(z-\Delta)^2}V_{-+}V_+V_{+-} + \sum_{k=2}^{\infty}\frac{V_{-+}V_+^kV_{+-}}{(z-\Delta)^{k+1}} \tag{6.34}$$

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

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 $A \otimes B \otimes C$ term:

$$\begin{aligned}\frac{1}{(z-\Delta)^2}(\kappa A+\lambda B)^2\mu C &= \underbrace{\frac{2\kappa\lambda\mu}{\Delta^2}A\otimes B\otimes C}_{(g_1)}+\underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)2\kappa\lambda\mu A\otimes B\otimes C}_{(g_2)} \\ &+ \underbrace{\frac{1}{(z-\Delta)^2}(\kappa^2+\lambda^2)\mu C}_{(g_3)}.\end{aligned} \tag{6.35}$$

Here $(g_1)$ combined with term $(a)$ in (6.34) gives $H_{\text{targ}}$. $(g_2)$ and $(g_3)$ are error terms. Now we further rearrange the error terms as the following. We combine term $(b)$ and $(e)$ to form $E_1$, term $(c)$ and $(g_3)$ to form $E_2$, term $(f)$ and the factors outside the bracket to be $E_3$. Rename $(g_2)$ to be $E_4$. Using the identity $(\kappa A+\lambda B)(A\otimes B)(\kappa A+\lambda B)=\text{sgn}(\alpha)(\kappa A+\lambda B)^2$ we combine term $(d)$ and $(h)$ along with the factors outside the bracket to be $E_5$. Rename $(i)$ to be $E_6$ and $(j)$ to be $E_7$. The rearranged self-energy expanision reads

$$\begin{aligned}\Sigma_-(z) &= \Bigg[\underbrace{H_{\text{else}}+\frac{2\kappa\lambda\mu}{\Delta^2}A\otimes B\otimes C}_{H_{\text{targ}}}+\underbrace{\left(\frac{1}{\Delta}+\frac{1}{z-\Delta}\right)(\kappa A+\lambda B)^2}_{E_1} \\ &+ \underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)(\kappa^2+\lambda^2)\mu C}_{E_2}+\underbrace{\frac{1}{(z-\Delta)^2}(\kappa A+\lambda B)H_{\text{else}}(\kappa A+\lambda B)}_{E_3} \\ &+ \underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)2\kappa\lambda\mu A\otimes B\otimes C}_{E_4}+\underbrace{\left(\frac{1}{(z-\Delta)^2}-\frac{1}{\Delta^2}\right)\frac{2\kappa\lambda}{\Delta}\text{sgn}(\alpha)(\kappa A+\lambda B)^2}_{E_5} \\ &- \underbrace{\frac{1}{(z-\Delta)^2}\cdot\frac{4\kappa^2\lambda^2}{\Delta^3}(\kappa A+\lambda B)^2}_{E_6}\Bigg]\otimes|0\rangle\langle 0|_w+\underbrace{\sum_{k=2}^{\infty}\frac{V_{-+}V_+^kV_{+-}}{(z-\Delta)^{k+1}}}_{E_7}.\end{aligned} \tag{6.36}$$

We bound the norm of each error term in the self energy expansion Eq. 6.36 by substituting the definitions of $\kappa$, $\lambda$ and $\mu$ in Eq. 6.32 and letting $z$ be the maximum value permitted by

Theorem 6.2.1 which is $\max z = |\alpha| + \epsilon + \|H_{\text{else}}\|$:

$$\|E_1\| \leq \frac{\max z \cdot 2^{4/3}\alpha^{2/3}\Delta^{2r-1}}{\Delta - \max z} = \Theta(\Delta^{2r-2}), \qquad \|E_2\| \leq \frac{(2\Delta - \max z)\max z}{(\Delta - \max z)^2} \cdot \alpha = \Theta(\Delta^{-1}) \tag{6.37}$$

$$\|E_3\| \leq \frac{2^{4/3}\alpha^{2/3}\Delta^{2r}\|H_{\text{else}}\|}{(\Delta - \max z)^2} = \Theta(\Delta^{2r-2}), \qquad \|E_4\| \leq \frac{(2\Delta - \max z)\max z}{(\Delta - \max z)^2} \cdot \alpha = \Theta(\Delta^{-1}) \tag{6.38}$$

$$\|E_5\| \leq \frac{(2\Delta - \max z)\max z}{(\Delta - \max z)^2} \cdot 2^{5/3}\alpha^{4/3}\Delta^{4r-3} = \Theta(\Delta^{4r-4}), \quad \|E_6\| \leq \frac{4\alpha^2\Delta^{6r-3}}{(\Delta - \max z)^2} = \Theta(\Delta^{6r-5}) \tag{6.39}$$

$$\begin{aligned} \|E_7\| &\leq \sum_{k=2}^{\infty} \left\| \frac{(\kappa A + \lambda B)\left(H_{\text{else}} + \mu C + \frac{2\kappa\lambda}{\Delta}\left(1 + \frac{2\kappa\lambda}{\Delta^2}\right) A \otimes B\right)^k (\kappa A + \lambda B)}{(\Delta - \max z)^{k+1}} \right\| \\ &\leq \frac{2^{4/3}\alpha^{2/3}\Delta^{2r}}{(\Delta - \max z)} \sum_{k=2}^{\infty} \frac{\left(\|H_{\text{else}}\| + 2^{-1/3}\alpha^{1/3}\Delta^{2-2r} + 2^{1/3}\alpha^{2/3}\Delta^{2r-1} + 2^{2/3}\alpha^{4/3}\Delta^{4r-3}\right)^k}{(\Delta - \max z)^k} \\ &= \Theta(\Delta^{\max\{1-2r, 6r-5, 10r-9\}}). \end{aligned} \tag{6.40}$$

Now the self energy expansion can be written as

$$\Sigma_-(z) = H_{\text{targ}} \otimes |0\rangle\langle 0|_w + \Theta(\Delta^{f(r)})$$

where the function $f(r) < 0$ determines the dominant power in $\Delta$ from $\|E_1\|$ through $\|E_6\|$:

$$f(r) = \max\{1 - 2r, 6r - 5\}, \quad \frac{1}{2} < r < 1. \tag{6.41}$$

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 $\epsilon$, $f(r)$ must take the minimum value. As is shown in Fig. 6.5b, 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 $\Delta$ in the assignments of $\kappa$, $\lambda$ and $\mu$ in Eq. 6.32 are optimal for the improved gadget construction. The optimal scaling of $\Theta(\epsilon^{-2})$ is also numerically confirmed in Fig. 6.4a. As one can see, the optimized slope $\mathrm{d}\log\Delta/\mathrm{d}\log\epsilon^{-1}$ is approximately 2 for small $\epsilon$.

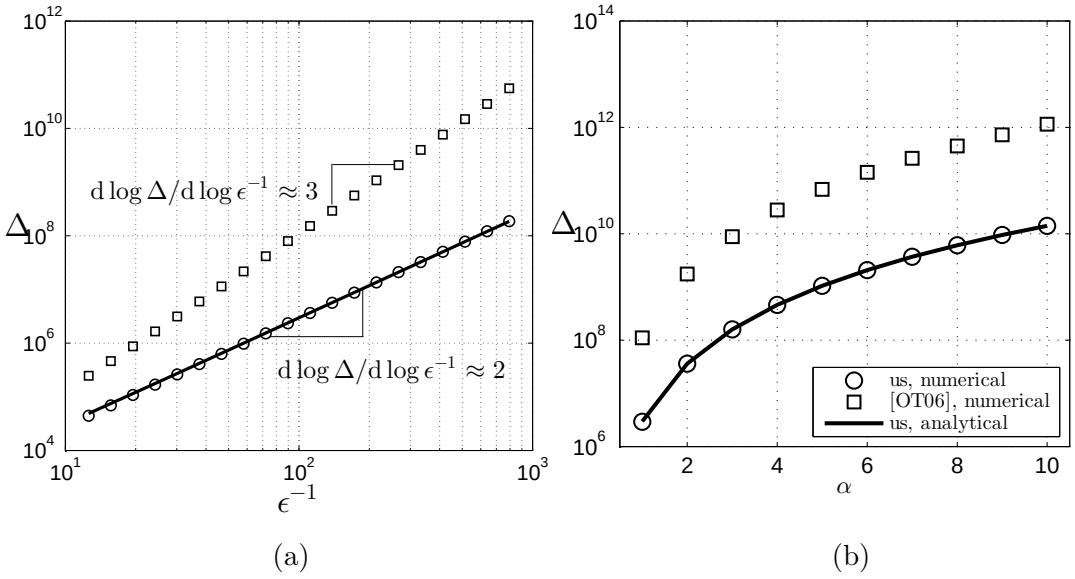



Figure 6.4: Comparison between our 3- to 2-body gadget with that of Oliveira and Terhal [216]. As $\Delta$ is not explicitly assigned as a function of $\alpha$, $\|H_{\text{else}}\|$ and $\epsilon$ in [216], we numerically find the optimal $\Delta$ values for their constructions (marked as "[OT06]"). (a) shows the scaling of the gap $\Delta$ as a function of error tolerance $\epsilon$. (b) shows the gap $\Delta$ as a function of the desired coupling $\alpha$. For the meanings of the labels in the legend, see Fig. 6.2. The fixed parameters in each subplots are: (a) $\|H_{\text{else}}\| = 0$, $\alpha = 1$. (b) $\epsilon = 0.01$, $\|H_{\text{else}}\| = 0$. Note that our constructions have improved the $\Delta$ scaling for the ranges of $\alpha$ and $\epsilon$ considered.

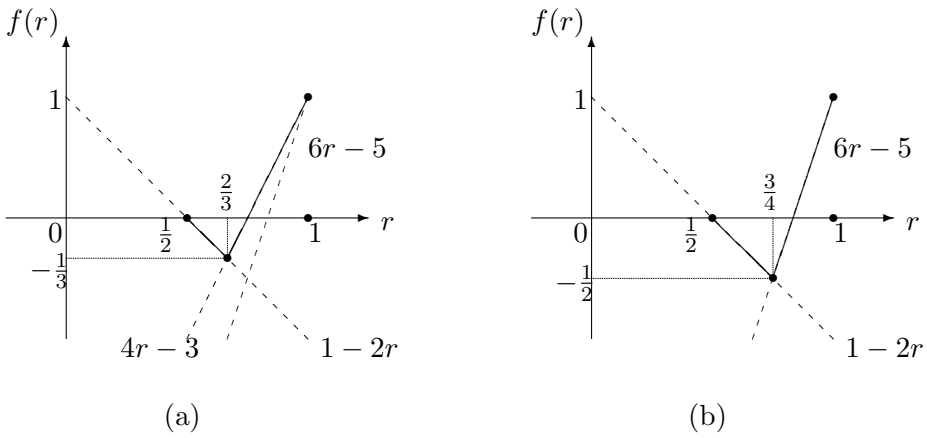



Figure 6.5: The function $f(r)$ shows the dominant power of $\Delta$ in the error terms in the perturbative expansion. (a) When the error term $E_4$ in Eq. 6.49, which contributes to the $4r-3$ component of $f(r)$ in Eq. 6.51, is not compensated in the original construction by Oliveira and Terhal, the dominant power of $\Delta$ 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 $\Delta = \Theta(\epsilon^{-2})$.

One natural question to ask next is whether it is possible to further improve the gap scaling as a function of $\epsilon$. This turns out to be difficult. Observe that the $6r-5$ component of $f(r)$ in Eq. 6.41 comes from $E_6$ and $E_7$ in Eq. 6.36. In $E_7$, the $\Theta(\Delta^{6r-5})$ contribution is attributed to the term $\frac{1}{\Delta}(\kappa A + \lambda B)^2$ in $V_1$ of Eq. 6.29, which is intended for compensating the $2^{\text{nd}}$ 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 $\Delta$ such that for given $\alpha$, $H_{\text{else}}$ and $\epsilon$, the spectral error between the effective Hamiltonian $H_{\text{eff}} = H_{\text{targ}} \otimes |0\rangle\langle 0|_w$ and $\tilde{H}_-$ is within $\epsilon$. This amounts to satisfying the condition of Theorem 6.2.1:

$$\|\Sigma_-(z) - H_{\text{eff}}\| \leq \epsilon. \tag{6.42}$$

Define the total error $E = \Sigma_-(z) - H_{\text{eff}} = E_1 + \cdots + E_7$. For convenience we also define $\eta = \|H_{\text{else}}\| + 2^{2/3}\alpha^{4/3}$ and $\xi = 2^{-1/3}\alpha^{1/3} + 2^{1/3}\alpha^{2/3}$. Then

$$\|E_7\| \leq \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}}{\Delta - \max z}\sum_{k=2}^{\infty}\frac{(\eta + \xi\Delta^{1/2})^k}{(\Delta - \max z)^k} = \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}}{\Delta - \max z - (\eta + \xi\Delta^{1/2})}\left(\frac{\eta + \xi\Delta^{1/2}}{\Delta - \max z}\right)^2. \tag{6.43}$$

The upper bound for $\|E\|$ is then found by summing over Eq. 6.37, 6.38, 6.39 and 6.43:

$$\begin{aligned}\|E\| \leq\ & \frac{\max z \cdot 2^{4/3}\alpha^{2/3}\Delta^{1/2}}{\Delta - \max z} + \frac{(2\Delta - \max z)\max z}{(\Delta - \max z)^2}\cdot 2^{4/3}\alpha^{3/2}\xi + \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}\eta}{(\Delta - \max z)^2} \\ +\ & \frac{2^{4/3}\alpha^{2/3}\Delta^{3/2}}{\Delta - \max z - (\eta + \xi\Delta^{1/2})}\left(\frac{\eta + \xi\Delta^{1/2}}{\Delta - \max z}\right)^2.\end{aligned} \tag{6.44}$$

By rearranging the terms in Eq. 6.44 we arrive at a simplified expression for the upper bound presented below. Requiring the upper bound of $\|E\|$ to be within $\epsilon$ gives

$$\|E\| \leq 2^{4/3}\alpha^{2/3}\frac{(\max z + \eta + \xi^2)\Delta^{1/2} + \xi(\max z + \eta)}{\Delta - \xi\Delta^{1/2} - (\max z + \eta)} \leq \epsilon. \tag{6.45}$$

Eq. 6.45 is a quadratic constraint with respect to $\Delta^{1/2}$. Solving the inequality gives the lower bound of $\Delta$ given in Eq. 6.30. Note here that $\Delta = \Theta(\epsilon^{-2})$, which improves over the previously assumed $\Delta = \Theta(\epsilon^{-3})$ in the literature [216, 165, 54]. This bound is shown in Fig. 6.4b as the "analytical lower bound". Comparison between the analytical lower bound and the numerically optimized gap in Fig. 6.4b indicates that the lower bound is relatively tight when $\|H_{\text{else}}\| = 0$.

*Comparison with Oliveira and Terhal [216].* Given operators $Q$, $R$ and $T$ acting on separate spaces $\mathcal{A}$, $\mathcal{B}$ and $\mathcal{C}$ respectively, the 3- to 2-body construction in [216, 165] approximates the target Hamiltonian $H_{\text{targ}} = H_{\text{else}} + Q \otimes R \otimes T$. In order to compare with their construction, however, we let $\alpha = \|Q\| \cdot \|R\| \cdot \|T\|$ and define $Q = \alpha^{1/3} A$, $R = \alpha^{1/3} B$ and $T = \alpha^{1/3} C$. Hence the target Hamiltonian $H_{\text{targ}} = H_{\text{else}} + \alpha A \otimes B \otimes C$ with $A$, $B$ and $C$ being unit-norm Hermitian operators. Introduce an ancilla qubit $w$ and apply the penalty Hamiltonian $H = \Delta |1\rangle\langle 1|_w$. In the construction by Oliveira and Terhal [216], the perturbation $V$ is defined as

$$V = H_{\text{else}} \otimes I_w + \mu C \otimes |1\rangle\langle 1|_w + (\kappa A + \lambda B) \otimes X_w + V_1' \tag{6.46}$$

where the compensation term $V_1'$ is

$$V_1' = \frac{1}{\Delta}(\kappa A + \lambda B)^2 - \frac{1}{\Delta^2}(\kappa^2 A^2 + \lambda^2 B^2)\mu C. \tag{6.47}$$

Comparing Eq. 6.47 with the expression for $V_1$ in Eq. 6.29, one observes that $V_1$ slightly improves over $V_1'$ by projecting 1-local terms to $\mathcal{L}_-$ so that $V$ will have less contribution to $V_+$, 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 $V$.

From the gadget construction shown in [216, Eq. 26], the equivalent choices of the

coefficients $\kappa$, $\lambda$ and $\mu$ are

$$\kappa = -\left(\frac{\alpha}{2}\right)^{1/3} \frac{1}{\sqrt{2}}\Delta^r, \quad \lambda = \left(\frac{\alpha}{2}\right)^{1/3} \frac{1}{\sqrt{2}}\Delta^r, \quad \mu = -\left(\frac{\alpha}{2}\right)^{1/3} \Delta^{2-2r} \tag{6.48}$$

where $r = 2/3$ in the constructions used in [216, 54]. 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 3$^{\text{rd}}$ order, following a similar procedure as in (6.34), we have

$$\begin{aligned}
\Sigma_-(z) = \Bigg[ & \underbrace{H_{\text{else}} + \frac{2\kappa\lambda\mu}{\Delta^2} A\otimes B\otimes C}_{H_{\text{targ}}} + \underbrace{\left(\frac{1}{\Delta} + \frac{1}{z-\Delta}\right)(\kappa A + \lambda B)^2}_{E_1} \\
& + \underbrace{\left(\frac{1}{(z-\Delta)^2} - \frac{1}{\Delta^2}\right)(\kappa^2 A^2 + \lambda^2 B^2)\mu C}_{E_2} + \underbrace{\frac{1}{(z-\Delta)^2}(\kappa A + \lambda B) H_{\text{else}} (\kappa A + \lambda B)}_{E_3} \\
& + \underbrace{\frac{1}{(z-\Delta)^2}\cdot\frac{1}{\Delta}(\kappa A + \lambda B)^4}_{E_4} - \underbrace{\frac{1}{(z-\Delta)^2}\cdot\frac{1}{\Delta^2}(\kappa^2 A^2 + \lambda^2 B^2)\mu(\kappa A + \lambda B)^2 \otimes C}_{E_5} \Bigg] \otimes |0\rangle\langle 0|_w \\
& + \underbrace{\sum_{k=2}^{\infty} \frac{V_{-+} V_+^k V_{+-}}{(z-\Delta)^{k+1}}}_{E_6}.
\end{aligned} \tag{6.49}$$

Similar to the derivation of Eq. 6.37, 6.38, and 6.39 by letting $z \mapsto \max z$, where $\max z = |\alpha| + \epsilon + \|H_{\text{else}}\|$ is the largest value of $z$ permitted by the Theorem 6.2.1, and using the triangle inequality to bound the norm, we can bound the norm of the error terms $E_1$ through $E_6$. For example,

$$\|E_1\| \leq \left(\frac{1}{\Delta - \max z} - \frac{1}{\Delta}\right) 2^2 \cdot \left(\frac{\alpha}{2}\right)^{2/3} \Delta^{2r} = \Theta(\Delta^{2r-2}).$$

Applying the same calculation to $E_2, E_3, \cdots$ we find that $\|E_2\| = \Theta(\Delta^{-1})$, $\|E_3\| = \Theta(\Delta^{2r-2})$, $\|E_4\| = \Theta(\Delta^{4r-3})$, $\|E_5\| = \Theta(\Delta^{4r-4})$. The norm of the high order terms $E_6$ can be bounded

as

$$\|E_6\| \leq \sum_{k=2}^{\infty} \frac{\|V_{-+}\| \cdot \|V_+\|^k \cdot \|V_{+-}\|}{(\Delta - \max(z))^{k+1}} \leq \frac{4\left(\frac{\alpha}{2}\right)^{1/3}\Delta^{2r}}{\Delta - \max(z)} \sum_{k=2}^{\infty} \left(\frac{\rho}{\Delta - \max(z)}\right)^k \tag{6.50}$$

$$= \frac{2^{4/3}\alpha^{2/3}\Delta^{2r}}{\Delta - \max(z) - \rho}\left(\frac{\rho}{\Delta - \max(z)}\right)^2 = \Theta(\Delta^{2r-1+2\max\{1-2r,2r-2\}}) = \Theta(\Delta^{\max\{1-2r,6r-5\}})$$

where $\rho = \|H_{\text{else}}\| + 2^{-1/3}\alpha^{1/3}\Delta^{2-2r} + 2^{1/3}\alpha^{2/3}\Delta^{2r-1}$. If we again write the self energy expansion Eq. 6.49 as

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

the function $f(r) < 0$, which determines the dominant power in $\Delta$ among $E_1$ through $E_6$, can be found as

$$f(r) = \max\{1-2r, 2r-2, 4r-3, 6r-5\}, \quad \frac{1}{2} < r < 1. \tag{6.51}$$

Similar to the discussion after Eq. 6.41, 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 Fig. 6.5a. Note that the $4r-3$ component in $f(r)$, Eq. 6.51, comes from the error term $E_4$ in Eq. 6.49. The idea for improving the gadget construction comes from the observation in Fig. 6.5a that when we add a term in $V$ to compensate for $E_4$, the dominant power of $\Delta$ in the perturbation series, $f(r)$, could admit a lower minimum as shown in Fig. 6.5b. 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 $O(\epsilon)$.

## 6.6 Creating 3-body gadget from local X

**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 [38], 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:

$$H_{ZZ} = \sum_i \delta_i X_i + \sum_i h_i Z_i + \sum_{i,j} J_{ij} Z_i Z_j. \tag{6.52}$$

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

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

$$V = H_{\text{else}} + \mu(Z_i + Z_j + Z_k) \otimes |1\rangle\langle 1|_w + \mu I \otimes X_w + V_{\text{comp}} \tag{6.53}$$

where $\mu = \left(\alpha\Delta^4/6\right)^{1/5}$ and $V_{\text{comp}}$ is

$$V_{\text{comp}} = \tag{6.54}$$
$$\frac{\mu^2}{\Delta}|0\rangle\langle 0|_w - \left(\frac{\mu^3}{\Delta^2} + 7\frac{\mu^5}{\Delta^4}\right)(Z_i + Z_j + Z_k) \otimes |0\rangle\langle 0|_w + \frac{\mu^4}{\Delta^3}\left(3I + 2Z_iZ_j + 2Z_iZ_k + 2Z_jZ_k\right).$$

To illustrate the basic idea of the $5^{\text{th}}$ order gadget, define subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$ in the usual way and define $P_-$ and $P_+$ as projectors into these respective subspaces. Then the second term in Eq. 6.53 with $\otimes|1\rangle\langle 1|_w$ contributes a linear combination $\mu Z_i + \mu Z_j + \mu Z_k$ to $V_+ = P_+ V P_+$. The third term in Eq. 6.53 induces a transition between $\mathcal{L}_-$ and $\mathcal{L}_+$ yet since it operates trivially on qubits 1-3, it only contributes a constant $\mu$ to the projections $V_{-+} = P_- V P_+$ and $V_{+-} = P_+ V P_-$. In the perturbative expansion, the $5^{\text{th}}$ order contains a term

$$\frac{V_{-+}V_+V_+V_+V_{+-}}{(z-\Delta)^4} = \frac{\mu^5(Z_i + Z_j + Z_k)^3}{(z-\Delta)^4} \tag{6.55}$$

due to the combined the contribution of the second and third term in Eq. 6.53. This yields a term proportional to $\alpha \cdot Z_i \otimes Z_j \otimes Z_k$ along with some 2-local error terms. These error terms,

combined with the unwanted terms that arise at $1^{\text{st}}$ through $4^{\text{th}}$ order perturbation, are compensated by $V_{\text{comp}}$. Note that terms at $6^{\text{th}}$ order and higher are $\Theta(\Delta^{-1/5})$. This means in order to satisfy the gadget theorem of Kempe *et al.* ([165, Theorem 3], or Theorem I.1) $\Delta$ needs to be $\Theta(\epsilon^{-5})$. This is the first perturbative gadget that simulates a 3-body target Hamiltonian using the Hamiltonian Eq. 6.52. By rotating the ancilla space, subdivision gadgets can also be implemented using this Hamiltonian: in the $X$ basis, $Z$ terms will induce a transition between the two energy levels of $X$. Therefore $Z_iZ_j$ coupling could be used for a perturbation of the form in Eq. 6.4 in the rotated basis. In principle using the transverse Ising model in Eq. 6.52, 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 $\Delta$ on the ancilla spin gives the unperturbed Hamiltonian $H = \Delta|1\rangle\langle 1|_w$. We then perturb the Hamiltonian $H$ using a perturbation $V$ described in (6.53). Using the same definitions of subspaces $\mathcal{L}_+$ and $\mathcal{L}_-$ as the previous 3-body gadget, the projections of $V$ into these subspaces can be written as

$$\begin{aligned}
V_+ &= \left\{ H_{\text{else}} + \mu(Z_1+Z_2+Z_3) + \frac{\mu^4}{\Delta^3}\big[3I + 2(Z_1Z_2 + Z_1Z_3 + Z_2Z_3)\big] \right\} \otimes |1\rangle\langle 1|_w \\
V_- &= \left\{ H_{\text{else}} + \frac{\mu^2}{\Delta} I - \frac{\mu^3}{\Delta^2}(Z_1+Z_2+Z_3)I + \frac{\mu^4}{\Delta^3}\big[3I + 2(Z_1Z_2 + Z_1Z_3 + Z_2Z_3)\big] \right. \\
&\quad \left. - \frac{7\mu^5}{\Delta^4}\big(Z_1+Z_2+Z_3\big) \right\} \otimes |0\rangle\langle 0|_w \\
V_{-+} &= \mu I \otimes |0\rangle\langle 1|_w, \quad V_{+-} = \mu I \otimes |1\rangle\langle 0|_w.
\end{aligned} \tag{6.56}$$

The low-lying spectrum of $\tilde{H}$ is approximated by the self energy expansion $\Sigma_-(z)$ below

with $z \in [-\max z, \max z]$ where $\max z = \|H_{\text{else}}\| + |\alpha| + \epsilon$. With the choice of $\mu$ above the expression of $V_+$ in Eq. 6.56 can be written as

$$V_+ = \Big(H_{\text{else}} + \mu(Z_1 + Z_2 + Z_3) + O(\Delta^{1/5})\Big) \otimes |1\rangle\langle 1|_w. \tag{6.57}$$

Because we are looking for the $5^{\text{th}}$ order term in the perturbation expansion that gives a term proportional to $Z_1 Z_2 Z_3$, expand the self energy in Eq. 6.3 up to $5^{\text{th}}$ order:

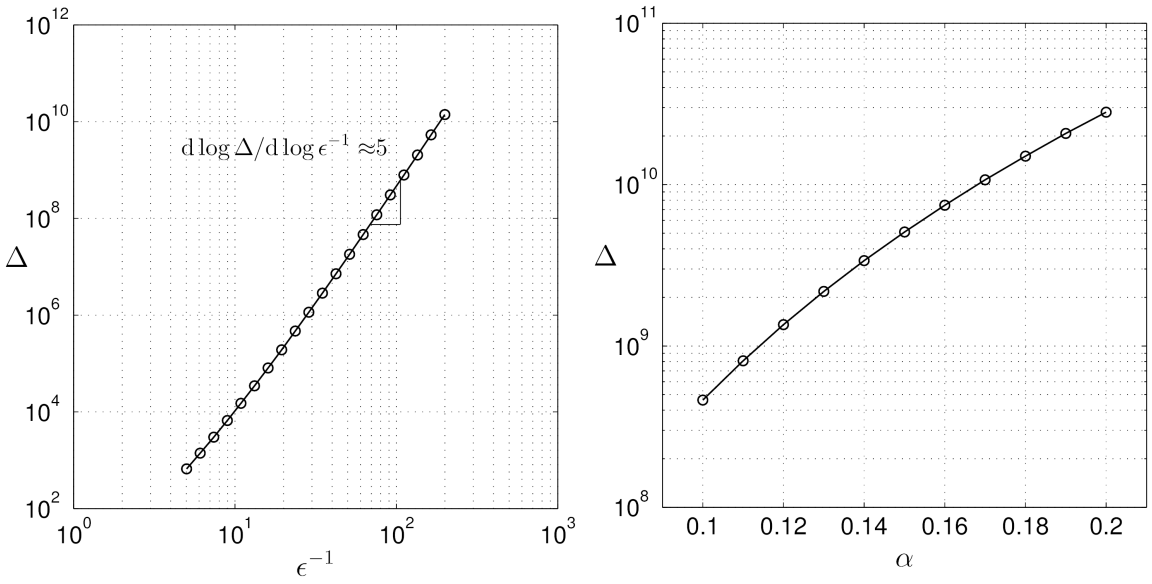
$$\begin{aligned}\Sigma_-(z) = V_- \otimes |0\rangle\langle 0|_w + \frac{V_{-+}V_{+-}}{z-\Delta} \otimes |0\rangle\langle 0|_w + \frac{V_{-+}V_+V_{+-}}{(z-\Delta)^2} \otimes |0\rangle\langle 0|_w \\ + \frac{V_{-+}V_+V_+V_{+-}}{(z-\Delta)^3} \otimes |0\rangle\langle 0|_w + \frac{V_{-+}V_+V_+V_+V_{+-}}{(z-\Delta)^4} \otimes |0\rangle\langle 0|_w + \sum_{k=4}^{\infty} \frac{V_{-+}V_+^k V_{+-}}{(z-\Delta)^{k+1}} \otimes |0\rangle\langle 0|_w.\end{aligned} \tag{6.58}$$

Using this simplification as well as the expressions for $V_-$, $V_{-+}$ and $V_{+-}$ in Eq. 6.56, the self energy expansion Eq. 6.58 up to $5^{\text{th}}$ order becomes

$$\begin{aligned}\Sigma_-(z) &= \underbrace{\left(H_{\text{else}} + \frac{6\mu^5}{\Delta^4} Z_1 Z_2 Z_3\right) \otimes |0\rangle\langle 0|_w}_{H_{\text{eff}}} + \underbrace{\left(\frac{1}{\Delta} + \frac{1}{z-\Delta}\right) \mu^2 I \otimes |0\rangle\langle 0|_w}_{E_1} \\ &+ \underbrace{\left(\frac{1}{(z-\Delta)^2} - \frac{1}{\Delta^2}\right) \mu^3 (Z_1 + Z_2 + Z_3) \otimes |0\rangle\langle 0|_w}_{E_2} \\ &+ \underbrace{\left(\frac{1}{\Delta^3} + \frac{1}{(z-\Delta)^3}\right) \cdot \mu^4 \cdot (Z_1 + Z_2 + Z_3)^2 \otimes |0\rangle\langle 0|_w}_{E_3} \\ &+ \underbrace{\left(\frac{1}{(z-\Delta)^4} - \frac{1}{\Delta^4}\right) 7\mu^5 (Z_1 + Z_2 + Z_3) \otimes |0\rangle\langle 0|_w}_{E_4} \\ &+ \underbrace{\frac{\mu^2}{(z-\Delta)^2} \cdot \frac{\mu^4}{\Delta^3} (Z_1 + Z_2 + Z_3)^2 \otimes |0\rangle\langle 0|_w}_{E_6} + \underbrace{\sum_{k=4}^{\infty} \frac{V_{-+}V_+^k V_{+-}}{(z-\Delta)^{k+1}} \otimes |0\rangle\langle 0|_w}_{E_7} \\ &+ O(\Delta^{-2/5}) + O(\|H_{\text{else}}\|\Delta^{-2/5}) + O(\|H_{\text{else}}\|^2 \Delta^{-7/5}) + O(\|H_{\text{else}}\|^3 \Delta^{-12/5}).\end{aligned} \tag{6.59}$$

Similar to what we have done in the previous sections, the norm of the error terms $E_1$ through $E_7$ can be bounded from above by letting $z \mapsto \max z$. Then we find that

$$\|\Sigma_-(z) - H_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \Theta(\Delta^{-1/5}) \tag{6.60}$$




(a) (b)

Figure 6.6: (a) The scaling of minimum $\Delta$ needed to ensure $\|\Sigma_-(z) - H_{\text{eff}}\| \leq \epsilon$ as a function of $\epsilon^{-1}$. Here we choose $\|H_{\text{else}}\| = 0$, $\alpha = 0.1$ and $\epsilon$ ranging from $10^{-0.7}$ to $10^{-2.3}$. The values of minimum $\Delta$ are numerically optimized. The slope of the line at large $\epsilon^{-1}$ is $4.97 \approx 5$, which provides evidence that with the assignments of $\mu = (\alpha\Delta^4/6)^{1/5}$, the optimal scaling of $\Delta$ is $\Theta(\epsilon^{-5})$. (b) The numerically optimized gap versus the desired coupling $\alpha$ in the target Hamiltonian. Here $\epsilon = 0.01$ and $\|H_{\text{else}}\| = 0$.

if we only consider the dominant dependence on $\Delta$ and regard $\|H_{\text{else}}\|$ as a given constant. To guarantee that $\|\Sigma_-(z) - H_{\text{targ}} \otimes |0\rangle\langle 0|_w\| \leq \epsilon$, we let the right hand side of Eq. 6.60 to be $\leq \epsilon$, which translates to $\Delta = \Theta(\epsilon^{-5})$.

This $\Theta(\epsilon^{-5})$ scaling is numerically illustrated (Fig. 6.6a). Although in principle the $5^{\text{th}}$ order gadget can be implemented on a Hamiltonian of form Eq. 6.52, for a small range of $\alpha$, the minimum $\Delta$ needed is already large (Fig. 6.6b), 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 $H_{\text{else}}$ is not necessarily diagonal.

## 6.7 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 [38], 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 [38], which is intended to simulate $Z_iX_j$ terms using Hamiltonians of the form

$$H_{ZZXX} = \sum_i \Delta_i X_i + \sum_i h_i Z_i + \sum_{i,j} J_{ij} Z_i Z_j + \sum_{i,j} K_{ij} X_i X_j, \tag{6.61}$$

is essentially a 3- to 2-body gadget with the target term $A \otimes B \otimes C$ being such that the operators $A$, $B$ and $C$ are $X$, $Z$ 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 $\Delta$ 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 Eq. 6.61. Otherwise it is not *a priori* clear how to realize YY terms using $H_{ZZXX}$ in Eq. 6.61. We will now present the first YY gadget which starts with a universal Hamiltonian of the form Eq. 6.61 and simulates the target Hamiltonian $H_{\text{targ}} = H_{\text{else}} + \alpha Y_i Y_j$. The basic idea is to use the identity $X_iZ_i = \iota Y_i$ where $\iota = \sqrt{-1}$ and induce a term of the form $X_iZ_iZ_jX_j = Y_iY_j$ at the $4^{\text{th}}$ order. Introduce ancilla qubit $w$ and apply a penalty $H = \Delta|1\rangle\langle 1|_w$. With a perturbation $V$ we could perform the same perturbative expansion as previously. Given that the $4^{\text{th}}$ order perturbation is $V_{-+}V_+V_+V_{+-}$ up to a scaling constant. we could let single $X_i$ and $X_j$ be coupled with $X_w$, which causes both $X_i$ and $X_j$ to appear in $V_{-+}$ and $V_{+-}$. Furthermore, we couple single $Z_i$ and $Z_j$ terms with $Z_w$. Then $\frac{1}{2}(I + Z_w)$ projects single $Z_i$ and $Z_j$ onto the $+$ subspace and causes them

to appear in $V_+$. For $H_{\rm targ} = H_{\rm else} + \alpha Y_1 Y_2$, the full expressions for the gadget Hamiltonian is the following: the penalty Hamiltonian $H = \Delta|1\rangle\langle 1|_w$ acts on the ancilla qubit. The perturbation $V = V_0 + V_1 + V_2$ where $V_0$, $V_1$, and $V_2$ are defined as

$$\begin{aligned} V_0 &= H_{\rm else} + \mu(Z_1 + Z_2) \otimes |1\rangle\langle 1|_w + \mu(X_1 - {\rm sgn}(\alpha) X_2) \otimes X_w \\ V_1 &= \frac{2\mu^2}{\Delta}(I \otimes |0\rangle\langle 0|_w + X_1 X_2) \\ V_2 &= -\frac{2\mu^4}{\Delta^3} Z_1 Z_2. \end{aligned} \tag{6.62}$$

with $\mu = (|\alpha|\Delta^3/4)^{1/4}$. For a specified error tolerance $\epsilon$, 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 $H_{\rm targ} \otimes |0\rangle\langle 0|_w$ up to error $\epsilon$.

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

$$H_{ZZYY} = \sum_i h_i X_i + \sum_i \Delta_i Z_i + \sum_{i,j} J_{ij} Z_i Z_j + \sum_{i,j} K_{ij} Y_i Y_j \tag{6.63}$$

and

$$H_{XXYY} = \sum_i h_i X_i + \sum_i \Delta_i Z_i + \sum_{i,j} J_{ij} X_i X_j + \sum_{i,j} K_{ij} Y_i Y_j \tag{6.64}$$

can be simulated using the Hamiltonian of the form in Eq. 6.61. Therefore using the Hamiltonian in Eq. 6.61 one can in principle simulate any finite-norm real valued Hamiltonian on qubits. Although by the QMA-completeness of $H_{ZZXX}$ one could already simulate such Hamiltonian via suitable embedding, our YY gadget provides a more direct alternative for the simulation.

**Analysis.** The results in [38] shows that Hamiltonians of the form in Eq. 6.61 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. Here we show that with a Hamiltonian of the

form in Eq. 6.61 we could simulate a target Hamiltonian $H_{\rm targ} = H_{\rm else} + \alpha Y_1 Y_2$. Introduce an ancilla $w$ and define the penalty Hamiltonian as $H = \Delta|1\rangle\langle 1|_w$. Let the perturbation $V = V_0 + V_1 + V_2$ be

$$\begin{aligned} V_0 &= H_{\rm else} + \kappa(Z_1 + Z_2) \otimes |1\rangle\langle 1|_w + \kappa(X_1 - {\rm sgn}(\alpha) X_2) \otimes X_w \\ V_1 &= 2\kappa^2 \Delta^{-1}[|0\rangle\langle 0|_w - {\rm sgn}(\alpha) X_1 X_2] \\ V_2 &= -4\kappa^4 \Delta^{-3} Z_1 Z_2. \end{aligned} \tag{6.65}$$

Then the gadget Hamiltonian $\tilde{H} = H + V$ is of the form in Eq. 6.61. Here we choose the parameter $\kappa = (|\alpha|\Delta^3/4)^{1/4}$. In order to show that the low lying spectrum of $\tilde{H}$ captures that of the target Hamiltonian, define $\mathcal{L}_- = {\rm span}\{|\psi\rangle \text{ such that } \tilde{H}|\psi\rangle = \lambda|\psi\rangle, \lambda < \Delta/2\}$ as the low energy subspace of $\tilde{H}$ and $\mathcal{L}_+ = I - \mathcal{L}_-$. Define $\Pi_-$ and $\Pi_+$ as the projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$ respectively.

With these notations in place, here we show that the spectrum of $\tilde{H}_- = \Pi_- \tilde{H} \Pi_-$ approximates the spectrum of $H_{\rm targ} \otimes |0\rangle\langle 0|_w$ with error $\epsilon$. To begin with, the projections of $V$ into the subspaces $\mathcal{L}_-$ and $\mathcal{L}_+$ can be written as

$$\begin{aligned} V_- &= \Big( H_{\rm else} + \underbrace{\frac{\kappa^2}{\Delta}(X_1 - {\rm sgn}(\alpha) X_2)^2}_{(a)} \underbrace{- \frac{4\kappa^4}{\Delta^3} Z_1 Z_2}_{(b)} \Big) \otimes |0\rangle\langle 0|_w \\ V_+ &= \Big( H_{\rm else} + \kappa(Z_1 + Z_2) - \frac{2\kappa^2}{\Delta} {\rm sgn}(\alpha) X_1 X_2 - \frac{4\kappa^4}{\Delta^3} Z_1 Z_2 \Big) \otimes |1\rangle\langle 1|_w \\ V_{-+} &= \kappa(X_1 - {\rm sgn}(\alpha) X_2) \otimes |0\rangle\langle 1|_w \\ V_{+-} &= \kappa(X_1 - {\rm sgn}(\alpha) X_2) \otimes |1\rangle\langle 0|_w \end{aligned} \tag{6.66}$$

Given the penalty Hamiltonian $H$, we have the operator valued resolvent $G(z) = (zI - H)^{-1}$ that satisfies $G_+(z) = \Pi_+ G(z) \Pi_+ = (z - \Delta)^{-1} |1\rangle\langle 1|_w$. Then the low lying sector of the gadget Hamiltonian $\tilde{H}$ can be approximated by the perturbative expansion Eq. 6.3. For

our purposes we will consider terms up to the $4^{\text{th}}$ order:

$$\Sigma_-(z) = V_- + \frac{1}{z-\Delta}V_{-+}V_{+-} + \frac{1}{(z-\Delta)^2}V_{-+}V_+V_{+-} + \frac{1}{(z-\Delta)^3}V_{-+}V_+V_+V_{+-} + \sum_{k=3}^{\infty}\frac{V_{-+}V_+^kV_{+-}}{(z-\Delta)^{k+1}}. \tag{6.67}$$

Now we explain the perturbative terms that arise at each order. The $1^{\text{st}}$ order is the same as $V_-$ in Eq. 6.66. The $2^{\text{nd}}$ order term gives

$$\frac{1}{z-\Delta}V_{-+}V_{+-} = \underbrace{\frac{1}{z-\Delta}\cdot\kappa^2(X_1 - \text{agn}(\alpha)X_2)^2}_{(c)}\otimes|0\rangle\langle 0|_w. \tag{6.68}$$

At the $3^{\text{rd}}$ order, we have

$$\begin{aligned}\frac{1}{(z-\Delta)^2}V_{-+}V_+V_{+-} &= \Bigg(\frac{1}{(z-\Delta)^2}\cdot\kappa^2(X_1-\text{agn}(\alpha)X_2)H_{\text{else}}(X_1-\text{sgn}(\alpha)X_2) \\ &+ \underbrace{\frac{1}{(z-\Delta)^2}\frac{4\kappa^4}{\Delta}(X_1X_2-\text{sgn}(\alpha)I)}_{(d)}\Bigg)\otimes|0\rangle\langle 0|_w + O(\Delta^{-1/4}).\end{aligned} \tag{6.69}$$

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

$$\begin{aligned}\frac{1}{(z-\Delta)^3}V_{-+}V_+V_+V_{+-} = \Bigg(&\underbrace{\frac{1}{(z-\Delta)^3}\cdot 2\kappa^4(X_1-\text{sgn}(\alpha)X_2)^2}_{(e)} - \underbrace{\frac{1}{(z-\Delta)^3}4\kappa^4Z_1Z_2}_{(f)} \\ &+ \frac{4\kappa^4\text{sgn}(\alpha)}{(z-\Delta)^3}Y_1Y_2\Bigg)\otimes|0\rangle\langle 0|_w + O(\|H_{\text{else}}\|\cdot\Delta^{-3/4}) + O(\|H_{\text{else}}\|^2\cdot\Delta^{-1/2})\end{aligned} \tag{6.70}$$

Note that with the choice of $\kappa = (|\alpha|\Delta^3/4)^{1/4}$, all terms of $5^{\text{th}}$ order and higher are of norm $O(\Delta^{-1/4})$. In the $1^{\text{st}}$ order through $4^{\text{th}}$ order perturbations the unwanted terms are labelled as $(a)$ through $(f)$ in Eqs. 6.66, 6.68, 6.69, and 6.70. 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

$$\Sigma_-(z) = (H_{\text{else}} + \alpha Y_1Y_2)\otimes|0\rangle\langle 0|_w + O(\Delta^{-1/4}). \tag{6.71}$$

Let $\Delta = \Theta(\epsilon^{-4})$, then by the Gadget Theorem (6.2.1), the low-lying sector of the gadget Hamiltonian $\tilde{H}_-$ captures the spectrum of $H_{\text{targ}} \otimes |0\rangle\langle 0|_w$ up to error $\epsilon$.

The fact that the gadget relies on $4^{\text{th}}$ 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.

## 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.

## Acknowledgements

We thank Andrew Landahl for helpful comments. JDB and YC completed parts of this study while visiting the Institute for Quantum Computing at the University of Waterloo. RB was supported by the United States Department of Defense. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government. JDB completed parts of this study while visiting the Qatar Energy and Environment Research Institute and would like to acknowledge the Foundational Questions Institute for support.

## 6.8 Appendix

### 6.8.1 Parallel 3- to 2-body gadget

**Summary.** In Sec. 6.4 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 $H_{\text{targ}} = H_{\text{else}} + \sum_{i=1}^{m} \alpha_i A_i \otimes B_i \otimes C_i$, 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 $H_{\text{else}}$ of the target Hamiltonian in reducing the next 3-body term. In this construction, $\Delta$ 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 $\Delta$ onto it. This poses additional challenges as the operator valued resolvent $G(z)$ now has multiple poles. Enumerating high order terms in the perturbation series requires consideration of the combinatorial properties of the bit flipping processes (Fig. 6.7).

If we apply the current construction [216, 54] of 3-body gadgets in parallel, which requires $\Delta = \Theta(\epsilon^{-3})$, it can be shown [54] 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 $A_i$, $B_i$ and $A_j$, $B_j$. 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 $H_{\text{targ}} = H_{\text{else}} + \alpha\sigma_1\sigma_2\cdots\sigma_k$ to 2-body [54]. In this paper 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 [54] for the range of $k$ concerned.

**Analysis.** In Sec. 6.4 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

$$H_{\text{targ}} = H_{\text{else}} + \sum_{i=1}^{m} \alpha_i A_i \otimes B_i \otimes C_i$$

where $H_{\text{else}}$ is a finite-norm Hamiltonian and all of $A_i$, $B_i$, $C_i$ are single-qubit Pauli operators acting on one of the $n$ qubits that $H_{\text{targ}}$ acts on. Here without loss of generality, we assume $A_i$, $B_i$ and $C_i$ 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, $A_i$, $B_i$ and $C_i$ act on three different qubits whose labels are in increasing order i.e. if we label the qubits with integers from 1 to $n$, $A_i$ acts on qubit $a_i$, $B_i$ acts on $b_i$, $C_i$ 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 $\Delta$ scales as $\Theta(\epsilon^{-2})$ instead of the common $\Theta(\epsilon^{-3})$ scaling assumed by the other 3-body constructions in the literature [165, 216, 54]. To reduce the $H_{\text{targ}}$ to 2-body, introduce $m$ qubits labelled as $u_1$, $u_2$, $\cdots$, $u_m$ and apply an energy penalty $\Delta$ onto the excited subspace of each qubit, as in the case of parallel subdivision gadgets presented previously. Then we have

$$H = \sum_{i=1}^{m} \Delta |1\rangle\langle 1|_{u_i} = \sum_{x \in \{0,1\}^m} h(x) \Delta |x\rangle\langle x|. \tag{6.72}$$

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

perturbation $V$ is defined as

$$V = H_{\text{else}} + \sum_{i=1}^{m} \mu_i C_i \otimes |1\rangle\langle 1|_{u_i} + \sum_{i=1}^{m} (\kappa_i A_i + \lambda_i B_i) \otimes X_{u_i} + V_1 + V_2 + V_3 \tag{6.73}$$

where $V_1$ is defined as

$$V_1 = \frac{1}{\Delta} \sum_{i=1}^{m} (\kappa_i A_i + \lambda_i B_i)^2 - \frac{1}{\Delta^2} \sum_{i=1}^{m} (\kappa_i^2 + \lambda_i^2) \mu_i C_i \tag{6.74}$$

and $V_2$ is defined as

$$V_2 = -\frac{1}{\Delta^3} \sum_{i=1}^{m} (\kappa_i A_i + \lambda_i B_i)^4. \tag{6.75}$$

$V_3$ will be explained later. Following the discussion in Sec. 6.5, the coefficients $\kappa_i$, $\lambda_i$ and $\mu_i$ are defined as

$$\kappa_i = \text{sgn}(\alpha_i) \left( \frac{|\alpha_i|}{2} \right)^{\frac{1}{3}} \Delta^{\frac{3}{4}}, \quad \lambda_i = \left( \frac{|\alpha_i|}{2} \right)^{\frac{1}{3}} \Delta^{\frac{3}{4}}, \quad \mu_i = \left( \frac{|\alpha_i|}{2} \right)^{\frac{1}{3}} \Delta^{\frac{1}{2}}. \tag{6.76}$$

However, as we will show in detail later in this section, a close examination of the perturbation expansion based on the $V$ in Eq. 6.73 shows that with assignments of $\kappa_i$, $\lambda_i$ and $\mu_i$ in Eq. 6.76 if $V$ has only $V_1$ and $V_2$ 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

$$V_3 = \sum_{i=1}^{m} \sum_{j=1, j \neq i}^{m} \bar{V}_{ij}$$

into $V$ and $\bar{V}_{ij}$ is the compensation term for cross-gadget contribution [1]. Before presenting

[1] As is shown by [54], for the gadget construction with the assignments of $\kappa_i$, $\lambda_i$ and $\mu_i$ all being $O(\Delta^{2/3})$, the cross-gadget contribution can be reduced by increasing $\Delta$, thus no cross-gadget compensation is needed. However, with our assignments of $\kappa_i$, $\lambda_i$ and $\mu_i$ in (6.76) there are cross-gadget error terms in the perturbative expansion that are of order $O(1)$, which cannot be reduced by increasing $\Delta$. 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.

the detailed form of $\bar{V}_{ij}$, let $s_1^{(i,j)} = s_{11}^{(i,j)} + s_{12}^{(i,j)}$ where

$$s_{11}^{(i,j)} = \begin{cases} 1 \quad \text{if} \begin{cases} [A_i, A_j] \neq 0 \\ [B_i, B_j] = 0 \end{cases} \text{or} \begin{cases} [B_i, B_j] \neq 0 \\ [A_i, A_j] = 0 \end{cases} \\ 0 \quad \text{otherwise} \end{cases} \tag{6.77}$$

$$s_{12}^{(i,j)} = \begin{cases} 1 \quad \text{if } [A_i, B_j] \neq 0 \text{ or } [B_i, A_j] \neq 0 \\ 0 \quad \text{otherwise} \end{cases} \tag{6.78}$$

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

$$s_2^{(i,j)} = \begin{cases} 1 \quad \text{if } [A_i, A_j] \neq 0 \text{ and } [B_i, B_j] \neq 0 \\ 0 \quad \text{otherwise.} \end{cases} \tag{6.79}$$

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

$$\bar{V}_{ij} \quad = \quad -s_1^{(i,j)} \cdot \frac{1}{\Delta^3}(\kappa_i \kappa_j)^2 I - s_2^{(i,j)} \left( \frac{2}{\Delta^3}(\kappa_i \kappa_j)^2 I - \frac{2}{\Delta^3} \kappa_i \kappa_j \lambda_i \lambda_j A_i A_j B_i B_j \right) \tag{6.80}$$

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 Eq. 6.80, although the term $A_i A_j B_i B_j$ 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 6.8.2. With the penalty Hamiltonian $H$ defined in Eq. 6.72, the operator-valued resolvent (or the Green's function) can be written as

$$G(z) = \sum_{x \in \{0,1\}^m} \frac{1}{z - h(x)\Delta} |x\rangle\langle x|. \tag{6.81}$$

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 $P_-$ and $P_+$ as the projectors onto $\mathcal{L}_-$ and $\mathcal{L}_+$. Then the projections of $V$

onto the subspaces can be written as

$$V_+ = \Bigg( H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)^2 - \frac{1}{\Delta^2}\sum_{i=1}^{m}(\kappa_i^2+\lambda_i^2)\mu_i C_i - \frac{1}{\Delta^3}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)^4$$
$$+ \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m}\bar{V}_{ij}\Bigg)\otimes P_+ + \sum_{i=1}^{m}\mu_i C_i \otimes P_+|1\rangle\langle 1|_{u_i}P_+ + \underbrace{\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)\otimes P_+ X_{u_i} P_+}_{V_f}$$

$$V_{-+} = \sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)\otimes P_- X_{u_i} P_+, \quad V_{+-} = \sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)\otimes P_+ X_{u_i} P_- \tag{6.82}$$

$$V_- = \Bigg( H_{\text{else}} + \frac{1}{\Delta}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)^2 - \frac{1}{\Delta^2}\sum_{i=1}^{m}(\kappa_i^2+\lambda_i^2)\mu_i C_i$$
$$- \frac{1}{\Delta^3}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)^4 + \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m}\bar{V}_{ij}\Bigg)\otimes P_-.$$

Here the $V_+$ projection is intentionally divided up into $V_f$ and $V_s$ components. $V_f$ is the component of $V_+$ that contributes to the perturbative expansion only when the perturbative term corresponds to flipping processes in the $\mathcal{L}_+$ subspace. $V_s$ 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 $G(z)$ onto $\mathcal{L}_+$ can be written as

$$G_+(z) = \sum_{x\neq 0\cdots 00}\frac{1}{z - h(x)\Delta}|x\rangle\langle x|. \tag{6.83}$$

We now explain the self energy expansion

$$\Sigma_-(z) = V_- + V_{-+}G_+V_{+-} + V_{-+}G_+V_+G_+V_{+-} + V_{-+}(G_+V_+)^2G_+V_{+-} + V_{-+}(G_+V_+)^3G_+V_{+-} + \cdots \tag{6.84}$$

in detail term by term. The 1st order term is simply $V_-$ from Equation Eq. 6.82. 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:

$$V_{-+}G_{+}V_{+-} = \frac{1}{z-\Delta}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)^2 \tag{6.85}$$

The 3$^{\text{rd}}$ 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 $V_+$ and then flipping the same qubit back. Therefore only the $V_f$ component in $V_+$ in Equation Eq. 6.82 will contribute to the perturbative expansion:

$$\begin{aligned} V_{-+}G_{+}V_{+}G_{+}V_{+-} &= \frac{1}{(z-\Delta)^2}\sum_{i=1}^{m}(\kappa_i A_i + \lambda_i B_i)\Big[H_{\text{else}} + \mu_i C_i + \frac{1}{\Delta}\sum_{j=1}^{m}(\kappa_j A_j + \lambda_j B_j)^2 \\ &+ \frac{1}{\Delta^2}\sum_{j=1}^{m}\Big[(\kappa_j^2+\lambda_j^2)\mu_j C_j - \frac{1}{\Delta^3}\sum_{j=1}^{m}(\kappa_j A_j + \lambda_j B_j)^4 + \sum_{j=1}^{m}\sum_{l=1,l\neq j}^{m}\bar{V}_{jl}\Big] \\ &\quad (\kappa_i A_i + \lambda_i B_i). \end{aligned} \tag{6.86}$$

The 4$^{\text{th}}$ order term is more involved. Here we consider two types of transition processes (for diagrammatic illustration refer to Fig. 6.7):

1. Starting from the all-zero state, flipping one of the qubits, flipping another qubit, then using the remaining $V_+$ and $V_{+-}$ to flip both qubits back one after the other (there are 2 different possible sequences, see Fig. 6.7a).

2. Starting from the all-zero state of the ancilla register, flipping one of the qubits, staying twice for the two $V_+$ components and finally flipping back the qubit during $V_{+-}$ (Fig. 6.7b).

Therefore in the transition processes of type (1), $V_+$ will only contribute its $V_f$ 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 Eq. 6.87 takes the form of a summation of two components. Because two qubits are flipped during the transition, $G_+$ 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), $V_+$ will only contribute its $V_s$ component to the $4^{\text{th}}$ order term since the states stay the same during both $V_+$ operators in the perturbative term. $G_+$ 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 in Eq. 6.87.

$$\begin{aligned}
V_{-+}(G_+V_+)^2G_+V_{+-} &= \frac{1}{(z-\Delta)^3}\sum_{i=1}^{m}(\kappa_iA_i+\lambda_iB_i)\Big[H_{\text{else}}+\mu_iC_i+\frac{1}{\Delta}\sum_{j=1}^{m}(\kappa_jA_j+\lambda_jB_j)^2 \\
&- \frac{1}{\Delta^2}\sum_{j=1}^{m}(\kappa_j^2+\lambda_j^2)\mu_jC_j-\frac{1}{\Delta^3}\sum_{j=1}^{m}(\kappa_jA_j+\lambda_jB_j)^4+\sum_{j=1}^{m}\sum_{l=1,l\neq j}^{m}\bar{V}_{jl}\Big]^2 \\
&\quad (\kappa_iA_i+\lambda_iB_i) \\
&+ \frac{1}{(z-\Delta)^2(z-2\Delta)}\sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m}\Big[(\kappa_iA_i+\lambda_iB_i)(\kappa_jA_j+\lambda_jB_j) \\
&\qquad\qquad (\kappa_iA_i+\lambda_iB_i)(\kappa_jA_j+\lambda_jB_j) \\
&+ (\kappa_iA_i+\lambda_iB_i)(\kappa_jA_j+\lambda_jB_j)(\kappa_jA_j+\lambda_jB_j)(\kappa_iA_i+\lambda_iB_i)\Big].
\end{aligned} \tag{6.87}$$

Although the $4^{\text{th}}$ order does not contain terms that are useful for simulating the 3-body target Hamiltonian, our assignments of $\kappa_i$, $\lambda_i$ and $\mu_i$ values in Eq. 6.76 imply that some of the terms at this order can be $\Theta(1)$. Indeed, the entire second term in Eq.6.87 is of order $\Theta(1)$ based on Eq. 6.76. Therefore it is necessary to study in detail what error terms arise at this order and how to compensate for them in the perturbation $V$. A detailed analysis on how to compensate the $\Theta(1)$ errors is presented in the Appendix 6.8.2. The $5^{\text{th}}$ order

and higher terms are errors that can be reduced by increasing $\Delta$:

$$\sum_{k=3}^{\infty} V_{-+}(G_+V_+)^k G_+ V_{+-}. \tag{6.88}$$

At first glance, with assignments of $\kappa_i$, $\lambda_i$ and $\mu_i$ in Eq. 6.76, it would appear that this error term is $\Theta(\Delta^{-1/4})$ since $\|V_{-+}\| = \Theta(\Delta^{3/4})$, $\|V_{+-}\| = \Theta(\Delta^{3/4})$, $\|V_+\| = \Theta(\Delta^{3/4})$ and $\|G_+\| = \Theta(\Delta^{-1})$,

$$\begin{aligned} \sum_{k=3}^{\infty} V_{-+}(G_+V_+)^k G_+ V_{+-} &\leq \sum_{k=3}^{\infty} \|V_{-+}\| \cdot \|G_+V_+\|^k \|G_+\| \cdot \|V_{+-}\| \\ &= \|V_{-+}(G_+V_+)^3 G_+ V_{+-}\| \sum_{k=0}^{\infty} \|G_+V_+\|^k \\ &= O(\Delta^{-1/4}) \end{aligned} \tag{6.89}$$

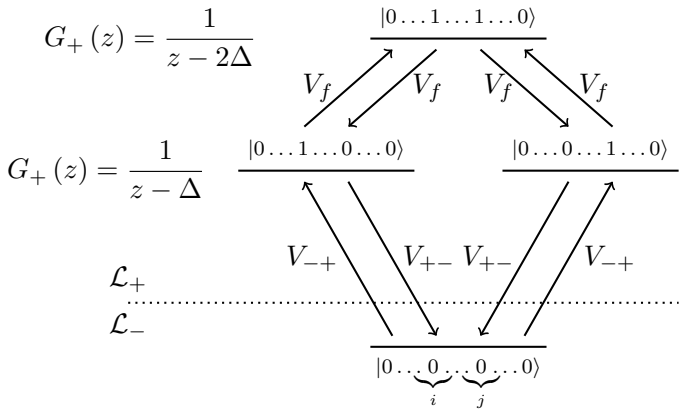
as $\sum_{k=0}^{\infty} \|G_+V_+\|^k = O(1)$. However, here we show that in fact this term in Eq. 6.88 is $\Theta(\Delta^{-1/2})$. Note that the entire term Eq. 6.88 consists of contributions from the transition processes where one starts with a transition from the all-zero state to a state $|x\rangle$ with $x \in \{0,1\}^m$ and $h(x) = 1$. If we focus on the perturbative term of order $k+2$:

$$V_{-+}(G_+V_+)^k G_+ V_{+-},$$

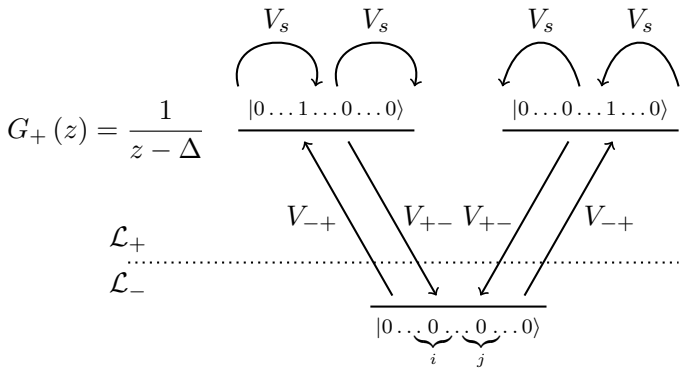
after $k$ steps. During every step one can choose to either flip one of the ancilla qubits or stay in the same state of the ancilla register, the state of the ancilla register will go back to a state $|y\rangle$ with $y \in \{0,1\}^m$ and $h(y) = 1$. Finally the $|1\rangle$ qubit in $|y\rangle$ is flipped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

$$K(k) = \begin{cases} \{k, k-2, \cdots, 2\} & \text{if } k \text{ is even} \\ \{k-1, k-3, \cdots, 2\} & \text{if } k \text{ is odd.} \end{cases} \tag{6.90}$$

For the term of order $k+2$, all the transition processes that contribute non-trivially to the term can be categorized into two types:




(a)




(b)

Figure 6.7: Diagrams illustrating the transitions that occur at 4th order. The two diagrams each represent a type of transition that occurs at 4th order. Each diagram is divided by a horizontal line where below the line is $\mathcal{L}_-$ space and above is $\mathcal{L}_+$ subspace. 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. $V_f$ and $V_s$ are both components of $V_+$. In fact $V_+ = V_f + V_s$ where $V_f$ is responsible for the flipping and $V_s$ contributes when the transition does not have flipping. At the left of each horizontal layer lies the expression for $G_+(z)$, 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 $V_s$ 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.

1. If $x = y$, the minimum number of flipping steps is 0. The contribution of all such processes to the $(k+2)$-th order perturbative term is bounded by [2]

$$\leq \quad m^{k_f} \cdot \binom{k}{k_f} \cdot \|V_f\|^{k_f} \cdot \|V_s\|^{k-k_f} \cdot \frac{\|V_{-+}\| \cdot \|V_{+-}\|}{(\Delta - \max(z))^{k+1}} \tag{6.91}$$

where the factor $m^{k_f}$ is the number of all possible ways of flipping $k_f$ times, each time one of the $m$ ancilla qubits. This serves as an upper bound for the number of transition processes that contribute non-trivially to the perturbative term. The factor $\binom{k}{k_f}$ describes the number of possible ways to choose which $(k - k_f)$ steps among the total $k$ steps involve the state of the ancilla register staying the same. $\|G_+\| \leq \frac{1}{\Delta - \max(z)}$ is used in the upper bound.

2. If $x \neq y$, the minimum number of flipping steps is 2. The contribution of all such processes to the $(k+2)$-th order perturbative term is bounded by

$$\leq \quad \binom{k}{k_f} \cdot \binom{k_f}{2} \cdot 2! \cdot \|V_f\|^{k_f} \|V_s\|^{k-k_f} \cdot m^{k_f - 2} \cdot \frac{\|V_{-+}\| \cdot \|V_{+-}\|}{(\Delta - \max(z))^{k+1}} \tag{6.92}$$

where the factor $\binom{k}{k_f}$ is the number of all possible ways to choose which $(k - k_f)$ steps among the $k$ steps should the state remain the same. $\binom{k_f}{2}$ is the number of possible ways to choose from the $k_f$ flipping steps the 2 minimum flips. 2! is for taking into account the ordering of the 2 flipping steps. $\|G_+\| \leq \frac{1}{\Delta - \max(z)}$ is used in the upper bound.

For a general $m$-qubit ancilla register, there are in total $m$ different cases of the first type of transition processes and $\binom{m}{2}$ different cases of the second type of transition processes.

[2] Here we use the notation $\mathsf{C}_m^n$ to represent the combinatorial number that is the number of ways to choose $n$ elements from a total of $m$ without distinguishing between different orderings.

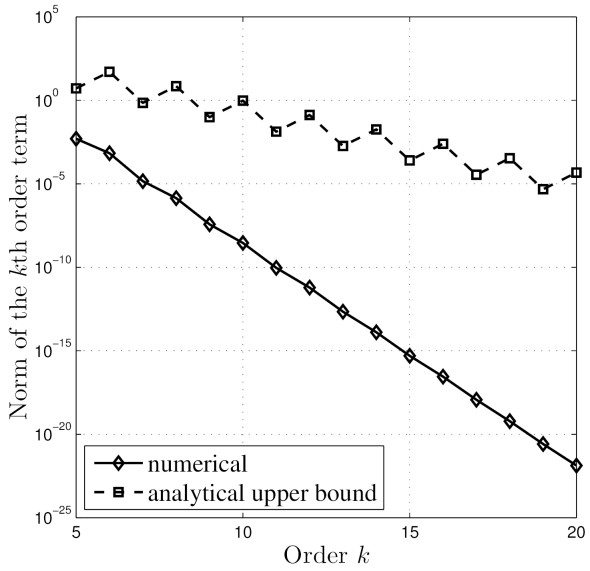



Figure 6.8: Numerical verification for the upper bound to the norm of the $(k+2)$-th order perturbative term in Eq.6.93. Here we use the parallel 3-body gadget for reducing $H_{\text{targ}} = 0.1X_1Z_2Z_3 - 0.2X_1X_2Z_3$ up to error $\epsilon = 0.01$. The gap in the gadget construction is numerically optimized. Here the calculation of the analytical upper bound uses the result in Eq.6.93. The calculation is then compared with the norm of the corresponding perturbative term numerically calculated according to the self-energy expansion.

Therefore we have the upper bound to the norm of the $(k+2)$-th term (Fig. 6.8)

$$\begin{aligned}
\|V_{-+}(G_+V_+)^k G_+V_{+-}\| &\leq m \sum_{k_f \in K(k)} m^{k_f} \binom{k}{k_f} \cdot \|V_f\|^{k_f} \cdot \|V_s\|^{k-k_f} \frac{\|V_{-+}\| \cdot \|V_{+-}\|}{(\Delta - \max(z))^{k+1}} \\
&+ \binom{m}{2} \sum_{k=3}^{\infty} \binom{k}{k_f} \cdot \binom{k_f}{2} \cdot 2! \cdot \|V_f\|^{k_f} \|V_s\|^{k-k_f} \cdot m^{k_f-2} \cdot \frac{\|V_{-+}\| \cdot \|V_{+-}\|}{(\Delta - \max(z))^{k+1}} \qquad (6.93) \\
&= \sum_{k_f \in K(k)} \left(m + \frac{m-1}{m}\right) 2^k \cdot \frac{\|V_{-+}\| \cdot (m\|V_f\|)^{k_f} \cdot \|V_s\|^{k-k_f} \cdot \|V_{+-}\|}{(\Delta - \max(z))^{k+1}} \\
&\leq \frac{\|V_{-+}\| \cdot \|V_{+-}\|}{\Delta - \max(z)} (m+1) \sum_{k=3}^{\infty} \left(\frac{\|V_s\|}{\Delta - \max(z)}\right)^k \sum_{k_f \in K(k)} \left(m \frac{\|V_f\|}{\|V_s\|}\right)^{k_f}.
\end{aligned}$$

Since $\| \sum_{i=1}^m \sum_{j=1, j\neq i}^m \bar{V}_{ij} \|$ is bounded by $\frac{1}{\Delta^3} \sum_{i=1}^m \sum_{j=1, j\neq i}^m 8(\kappa_i \kappa_j)^2 I$, from Eq. 6.82 we

see that

$$\begin{aligned}\|V_s\| &\leq \|H_{\text{else}}\| + 2^{-1/3}\Delta^{1/2}\sum_{i=1}^{m}|\alpha_i|^{1/3} + 2^{4/3}\Delta^{1/2}\sum_{i=1}^{m}|\alpha_i|^{2/3} + \sum_{i=1}^{m}|\alpha_i| \\ &\quad +2^{8/3}\sum_{i=1}^{m}|\alpha_i|^{4/3} + \sum_{i=1}^{m}\sum_{j=1,j\neq i}^{m} 8\cdot 2^{-4/3}|\alpha_i|^{2/3}|\alpha_j|^{2/3} \equiv v_s \\ \|V_f\| &\leq 2^{2/3}\Delta^{3/4}\sum_{i=1}^{m}|\alpha_i|^{1/3} \equiv v_f.\end{aligned} \tag{6.94}$$

With bounds of $\|V_s\|$ and $\|V_f\|$ in Eq.6.82, the summation in Equation Eq. 6.93 can be written as

$$\begin{aligned}&\|\sum_{k=3}^{\infty} V_{-+}(G_+V_+)^k G_+V_{+-}\| \leq \frac{\|V_{-+}\|\cdot\|V_{+-}\|}{\Delta-\max(z)}(m+1) \\ &\Big[\sum_{r=1}^{\infty}\left(\frac{2v_s}{\Delta-\max(z)}\right)^{2r+1}\sum_{s=1}^{r}\left(m\frac{v_f}{v_s}\right)^{2s} + \sum_{r=2}^{\infty}\left(\frac{2v_s}{\Delta-\max(z)}\right)^{2r}\sum_{s=1}^{r}\left(m\frac{v_f}{v_s}\right)^{2s}\Big].\end{aligned} \tag{6.95}$$

To guarantee convergence of the summation in Eq.6.95 we require that $\Delta$ satisfies

$$\frac{2mv_f}{\Delta-\max(z)} < 1 \tag{6.96}$$

$$m\left(\frac{v_f}{v_s}\right) > 1, \tag{6.97}$$

both of which are in general satisfied. The summation in Eq. 6.95 can then be written as

$$\begin{aligned}&\|\sum_{k=3}^{\infty} V_{-+}(G_+V_+)^k G_+V_{+-}\| \leq \frac{\|V_{-+}\|\cdot\|V_{+-}\|}{\Delta-\max(z)}\cdot\frac{\left(m\frac{v_f}{v_s}\right)^2}{\left(m\frac{v_f}{v_s}\right)^2-1} \\ &\frac{\left(\frac{2mv_f}{\Delta-\max(z)}\right)^2}{1-\left(\frac{2mv_f}{\Delta-\max(z)}\right)^2}(m+1)\left[\left(\frac{2mv_f}{\Delta-\max(z)}\right)^2 + \frac{2v_s}{\Delta-\max(z)}\right] = \Theta(\Delta^{-1/2}),\end{aligned} \tag{6.98}$$

which shows that the high order terms are $\Theta(\Delta^{-1/2})$. This is tighter than the crude bound $\Theta(\Delta^{-1/4})$ shown in Eq. 6.89. The self-energy expansion Eq. 6.84 then satisfies

$$\|\Sigma_-(z) - H_{\text{targ}}\otimes P_-\| \leq \Theta(\Delta^{-1/2}) \tag{6.99}$$

which indicates that $\Delta = \Theta(\epsilon^{-2})$ is sufficient for the parallel 3-body gadget to capture the entire spectrum of $H_{\text{targ}}\otimes P_-$ up to error $\epsilon$.

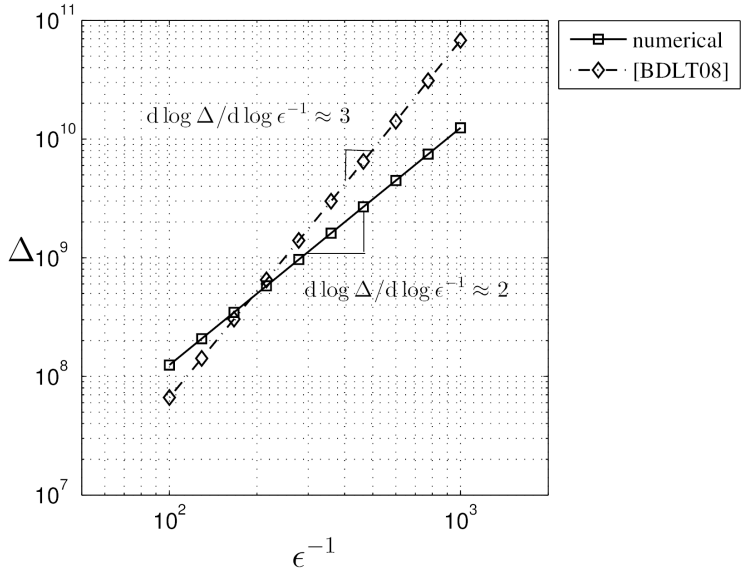



Figure 6.9: Scaling of the spectral gap $\Delta$ as a function of error $\epsilon$ for the parallel 3-body example that is intended to reduce the target Hamiltonian $H_{\mathrm{targ}} = Z_1Z_2Z_3 - X_1X_2X_3$ to 2-body. Here $\epsilon = 0.01$. We show both numerically optimized values ("numerical") in our construction and the construction in [54], which is referred to as "[Bravyi2008]".

We have used numerics to verify the $\Theta(\epsilon^{-2})$ scaling, as shown in Fig. 6.8. Furthermore, for a range of specified $\epsilon$, the minimum $\Delta$ needed for the spectral error between the gadget Hamiltonian and the target Hamiltonian is numerically found. In the optimized cases, the slope $\mathrm{d}\log\Delta/\mathrm{d}\log\epsilon^{-1}$ for the construction in [54] is approximately 3, showing that $\Delta = \Theta(\epsilon^{-3})$ is the optimal scaling for the construction in [54]. For our construction both the analytical bound and the optimized $\Delta$ scale as $\Theta(\epsilon^{-2})$ (see Fig. 6.9).

### 6.8.2 Compensation for the 4-local error terms in parallel 3- to 2-body gadget

Continuing the discussion in Appendix 6.8.1, here we deal with $\Theta(1)$ error terms that arise in the $3^{\mathrm{rd}}$ and $4^{\mathrm{th}}$ order perturbative expansion when $V$ in Eq. 6.73 is without $V_3$ and in so doing explain the construction of $\bar{V}_{ij}$ in Eq. 6.80. From the previous description of

the $3^{\text{rd}}$ and $4^{\text{th}}$ order terms, for each pair of terms $(i)$ and $(j)$ where $i$ and $j$ are integers between 1 and $m$, let

$$\begin{aligned} M_1 &= (\kappa_i A_i + \lambda_i B_i)(\kappa_j A_j + \lambda_j B_j) \\ M_2 &= (\kappa_j A_j + \lambda_j B_j)(\kappa_i A_i + \lambda_i B_i) \end{aligned}$$

and then the $\Theta(1)$ error term arising from the $3^{\text{rd}}$ and $4^{\text{th}}$ order perturbative expansion can be written as

$$\frac{1}{(z-\Delta)^2}\left[\frac{1}{z-2\Delta}(M_1^2+M_2^2)+\left(\frac{1}{\Delta}+\frac{1}{z-2\Delta}\right)(M_1M_2+M_2M_1)\right]. \tag{6.100}$$

Based on the number of non-commuting pairs among $A_i$, $A_j$, $B_i$ and $B_j$, all possible cases can be enumerated as the following:

$$\begin{aligned} &\text{case 0:} && [A_i,A_j]=0,[B_i,B_j]=0,[A_i,B_j]=0,[B_i,A_j]=0 \\ &\text{case 1:} && 1.1: \quad [A_i,A_j]=0,[B_i,B_j]=0,[A_j,B_i]\neq 0 \\ & && 1.2: \quad [A_i,A_j]=0,[B_i,B_j]=0,[A_i,B_j]\neq 0 \\ & && 1.3: \quad [A_i,A_j]=0,[B_i,B_j]\neq 0 \\ & && 1.4: \quad [A_i,A_j]\neq 0,[B_i,B_j]=0 \\ &\text{case 2:} && [A_i,A_j]\neq 0,[B_i,B_j]\neq 0. \end{aligned} \tag{6.101}$$

In case 0, clearly $M_1 = M_2$. Then the $\Theta(1)$ error becomes

$$\frac{1}{(z-\Delta)^2}\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2M_1^2 = \Theta(\Delta^{-1})$$

which does not need any compensation. In case 1, for example in the subcase 1.1, $A_j$ does not commute with $B_i$. Then $M_1$ and $M_2$ can be written as

$$\begin{aligned} M_1 &= K + \kappa_j\lambda_i B_i A_j \\ M_2 &= K + \kappa_j\lambda_i A_j B_i \end{aligned}$$

where $K$ contains the rest of the terms in $M_1$ and $M_2$. Furthermore,

$$M_1^2 + M_2^2 = 2K^2 - 2(\kappa_j\lambda_i)^2 I$$

$$M_1 M_2 + M_2 M_1 = 2K^2 + 2(\kappa_j\lambda_i)^2 I.$$

Hence the $\Theta(1)$ term in this case becomes

$$\frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta} + \frac{2}{z-2\Delta}\right) 2K^2 + \frac{1}{\Delta}\cdot 2(\kappa_j\lambda_i)^2 I\right] \tag{6.102}$$

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 $R = \kappa_i\lambda_j A_i B_j + \lambda_i\kappa_j B_i A_j$ and $T = \kappa_i\kappa_j A_i A_i + \lambda_i\lambda_j B_i B_i$. Then

$$M_1^2 + M_2^2 = 2(R^2 + T^2)$$

$$M_1 M_2 + M_2 M_1 = 2(R^2 - T^2).$$

The $\Theta(1)$ error terms in the 3rd and 4th order perturbative expansion becomes

$$\frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta} + \frac{2}{z-2\Delta}\right)\cdot 2R^2 - \frac{1}{\Delta}\cdot 2T^2\right] \tag{6.103}$$

where the first term is $\Theta(\Delta^{-1})$ and hence needs no compensation. The second term is $\Theta(1)$. Define

$$s_0^{(i,j)} = \begin{cases} 1 & \text{if case 0} \\ \\ 0 & \text{Otherwise} \end{cases} \tag{6.104}$$

With the definitions of $s_1^{(i,j)}$ and $s_2^{(i,j)}$ in Eq. 6.77, Eq. 6.78 and Eq. 6.79, the contribution of the $i$-th and the $j$-th target terms to the $\Theta(1)$ error in the perturbative expansion $\Sigma_-(z)$

becomes

$$\begin{aligned} & s_0^{(i,j)} \cdot \frac{1}{(z-\Delta)^2}\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2(\kappa_i A_i+\lambda_i B_i)^2(\kappa_j A_j+\lambda_j B_j)^2 \\ + \quad & s_1^{(i,j)} \cdot \frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2K_{ij}^2+\frac{1}{\Delta}\cdot 2(\kappa_i\kappa_j)^2 I\right] \\ + \quad & s_2^{(i,j)} \cdot \frac{1}{(z-\Delta)^2}\left[\left(\frac{1}{\Delta}+\frac{2}{z-2\Delta}\right)\cdot 2R_{ij}^2+\frac{1}{\Delta}\cdot 2\{[(\kappa_i\kappa_j)^2+(\lambda_i\lambda_j)^2]I \right. \\ & \left. -2\kappa_i\kappa_j\lambda_i\lambda_j A_i A_j B_i B_j\}\right]. \end{aligned} \tag{6.105}$$

The term proportional to $s_0^{(i,j)}$ in Eq. 6.105 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 $\bar{V}_{ij}$ in Eq. 6.80 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 $\bar{V}_{ij}$ in Eq. 6.80 that is proportional to $s_2^{(i,j)}$.

Now we deal with generating the 4-local term in $\bar{V}_{ij}$. Introduce an ancilla $u_{ij}$ and construct a gadget $\tilde{H}_{ij} = H_{ij} + V_{ij}$ such that $H_{ij} = \Delta|1\rangle\langle 1|_{u_{ij}}$ and the perturbation $V_{ij}$ becomes

$$V_{ij} = (\kappa_i A_i + \lambda_j B_j)\otimes X_{u_{ij}} + (\kappa_j A_j + \lambda_i B_i)\otimes |1\rangle\langle 1|_{u_{ij}} + V'_{ij} \tag{6.106}$$

where $V'_{ij}$ is defined as

$$V'_{ij} = \frac{1}{\Delta}(\kappa_i A_i + \lambda_j B_j)^2 + \frac{1}{\Delta^3}\left[(\kappa_j^2+\lambda_i^2)(\kappa_i A_i+\lambda_j B_j)^2 - 2\kappa_j\lambda_i(\kappa_j^2+\lambda_j^2)A_j B_i\right] \tag{6.107}$$

The self-energy expansion $\Sigma_-(z)$ is now

$$\Sigma_-(z) = \frac{1}{(z-\Delta)^3} 4\kappa_i\kappa_j\lambda_i\lambda_j A_i A_j B_i B_j + O(\Delta^{-1/2})$$

which is $O(\Delta^{-1/2})$ close to the 4-local compensation term in $\bar{V}_{ij}$. We apply the the gadget $\tilde{H}_{ij}$ for every pair of qubits with $s_2^{(i,j)} = 1$. The cross-gadget contribution between the $\tilde{H}_{ij}$ gadgets as well as those cross-gadget contribution between $\tilde{H}_{ij}$ gadgets and gadgets based on ancilla qubits $u_1$ through $u_m$ both belong to the case 1 of the Eq. 6.101 and hence are easy to deal with using 2-body terms.

## Chapter 7

# Adiabatic Quantum Simulation of Quantum Chemistry

Apart from minor modifications, this chapter originally appeared as [14]:

"Adiabatic Quantum Simulation of Quantum Chemistry". Ryan Babbush, Peter Love and Alán Aspuru-Guzik. *Scientific Reports*. Volume 4, Number 6603: 1-11. 2014.

## Abstract

We show how to apply the quantum adiabatic algorithm directly to the quantum computation of molecular properties. We describe a procedure to map electronic structure Hamiltonians to 2-body qubit Hamiltonians with a small set of physically realizable couplings. By combining the Bravyi-Kitaev construction to map fermions to qubits with perturbative gadgets to reduce the Hamiltonian to 2-body, we obtain precision requirements on the coupling strengths and a number of ancilla qubits that scale polynomially in the problem size. Hence our mapping is efficient. The required set of controllable interactions includes only two types of interaction beyond the Ising interactions required to apply the quantum adiabatic algorithm to combinatorial optimization problems. Our mapping may also be of interest to chemists directly as it defines a dictionary from electronic structure to spin Hamiltonians with physical interactions.

## 7.1 Introduction

The ability to make exact quantum chemical calculations on nontrivial systems would revolutionize chemistry. While seemingly intractable for classical algorithms, quantum computers can efficiently perform such computations. There has been substantial interest in quantum algorithms for quantum chemistry involving a combination of Trotterization and phase estimation [11, 285, 281, 140, 224, 195]. However, we are still technologically far from when such gate-model approaches are experimentally feasible for practical chemistry problems. Here, we propose a radically different approach based on the quantum adiabatic algorithm. In this rapidly advancing paradigm of quantum computation, there is no need for Trotterization, phase estimation or logic gates. More generally, we show the first scalable quantum simulation scheme for fermionic systems using adiabatic quantum computing.

Adiabatic quantum computing works by changing the Hamiltonian of a controllable quantum system from an initial Hamiltonian whose ground state is easy to prepare into a Hamiltonian whose ground state encodes the solution of a computationally interesting problem [104, 103]. The speed of this algorithm is determined by the adiabatic theorem of quantum mechanics which states that an eigenstate remains at the same position in the eigenspectrum if a perturbation acts on the system sufficiently slowly [48, 104, 46]. Simply embedding a computational problem in a Hamiltonian suitable for AQC does not ensure an efficient solution. The required runtime for the adiabatic evolution depends on the energy gap between the ground state and first excited state at the smallest avoided crossing [104].

AQC has been applied to classical optimization problems that lie in the complexity class NP [114]. For example, studies have been performed on satisfiability [148, 74, 205], Exact Cover [104, 103], 3-regular 3-XORSAT and 3-regular Max-Cut [105], random instances of classical Ising spin glasses [43, 44, 233], protein folding [221, 18] and machine learning [13,

86, 210]. AQC has also been applied to structured and unstructured search [231, 232], search engine ranking [115] and artificial intelligence problems arising in space exploration [253]. Many of these applications follow naturally from the NP-Completeness of determining the ground state energy of classical Ising spin glasses [23]. This creates an equivalence between a large set of computational problems (the class NP) and a set of models in classical physics (classical Ising models with random coupling strengths). The advent of AQC provides a powerful motivation to study the detailed implications of this mapping. In general, we do not expect that quantum computing, including AQC, can provide efficient solutions to NP-Complete problems *in the worst case* [32]. However, there may exist sets of instances of some NP-Complete problems for which AQC can find the ground state efficiently, but which defy efficient classical solution by any means. If this is the case then AQC is certainly of considerable scientific interest, and likely of great industrial importance.

The potential value of a positive answer to this conjecture has motivated a commercial effort to construct an adiabatic quantum computer [132, 134, 133, 136, 173, 153, 28]. Currently, these experimental implementations of AQC are not strictly confined to the ground state at zero temperature but have considerable thermal mixing of higher lying states. Such intermediate implementations are referred to as quantum annealing devices. Quantum annealing machines with up to 509 qubits have been commercially manufactured by *D-Wave Systems* [29, 154, 87]. They are currently the subject of serious scientific investigation to determine whether their operation depends significantly on their quantum properties, and if so, whether it provides a speedup for any class of instances [154, 225, 39, 44, 278, 254, 233, 174].

Quantum computers have been rigorously proved to provide an algorithmic advantage over the best known classical approaches for a small set of problems [251, 69, 124, 256]. Adiabatic quantum computation applied to classical Ising Hamiltonians (equivalently, all

problems in NP) also gives an approach to a very large class of problems where the advantage (if any) is currently unknown. The construction of medium scale (500 qubit) quantum annealing machines provides a hardware platform where the properties of AQC can be investigated experimentally. Such investigations have already been performed for many problems. At present, optimized codes on classical hardware can find the ground state of many instances in comparable time to the D-Wave device [43]. However, even if no interesting set of instances is found on which quantum annealing on the classical Ising model outperforms classical approaches, the hardware constructed to date represents an important step towards the construction of large scale quantum information technology. If quantum annealing of the classical Ising model is the first step, what is the natural next step?

Quantum simulation has provided a rich set of questions and methods in quantum computation since Feynman's suggestion that quantum devices would be best suited to computation of quantum properties [107]. This observation has been fleshed out through early work on specific systems [197, 289, 1, 182, 42, 298] and through quantum algorithms for computation of eigenvalues, dynamics and other properties [2, 33, 162, 287, 280, 237]. Recently, there have been many proposals for the simulation of quantum lattice models using trapped ions, trapped atoms and photonic systems [282, 128, 78, 141]. There has been rapid experimental progress in the quantum simulation of a number of systems [252, 120, 179, 109, 152, 188, 228]. A natural target for these simulations is the phase diagram of the Fermi-Hubbard model - believed to inform our understanding of high-$T_c$ superconductivity. For this reason many of these approaches are aimed at simulating systems of interacting fermions.

Lattice systems are a natural target for trapped ion and atom quantum simulators, with

the trapping mechanism taking the place of the crystal lattice and interactions restricted to neighbors on the lattice. However, quantum chemistry applied to molecular systems is perhaps the broadest class of problems on which quantum simulation of interacting fermions could have an impact. Finding the energy of electrons interacting in the Coulomb potential of a set of fixed nuclei of an atom or molecule defines the electronic structure problem. This problem appears to be hard for classical computers because the cost of directly solving for the eigenvalues of the exact electronic Hamiltonian grows exponentially with the problem size. In spite of much progress over the last 60 years developing approximate classical algorithms for this problem, exact calculations remain out of reach for many systems of interest. Figure 7.1 shows several of the proposals for the efficient quantum simulation of chemical Hamiltonians.

One may divide quantum simulation algorithms into two classes: those that address statics and compute ground state properties, and those that address dynamics, and simulate time evolution of the wavefunction. It is clear that the simulation of time evolution is exponentially more efficient on quantum computers, with significant implications for the simulation of chemically reactive scattering, in particular [163]. The computation of ground state properties naturally requires preparation of the ground state. This can be done adiabatically [11, 272], or by preparation of an ansatz for the ground state [223]. Adiabatic preparation of the ground state within a gate model simulation requires time evolution of the wavefunction, which is efficient. However, the length of time for which one must evolve is determined, as for all adiabatic algorithms, by the minimum energy gap between ground and first excited states along the adiabatic path. This is unknown in general. Similarly, a successful ansatz state must have significant overlap with the true ground state, and guarantees of this are unavailable in general.

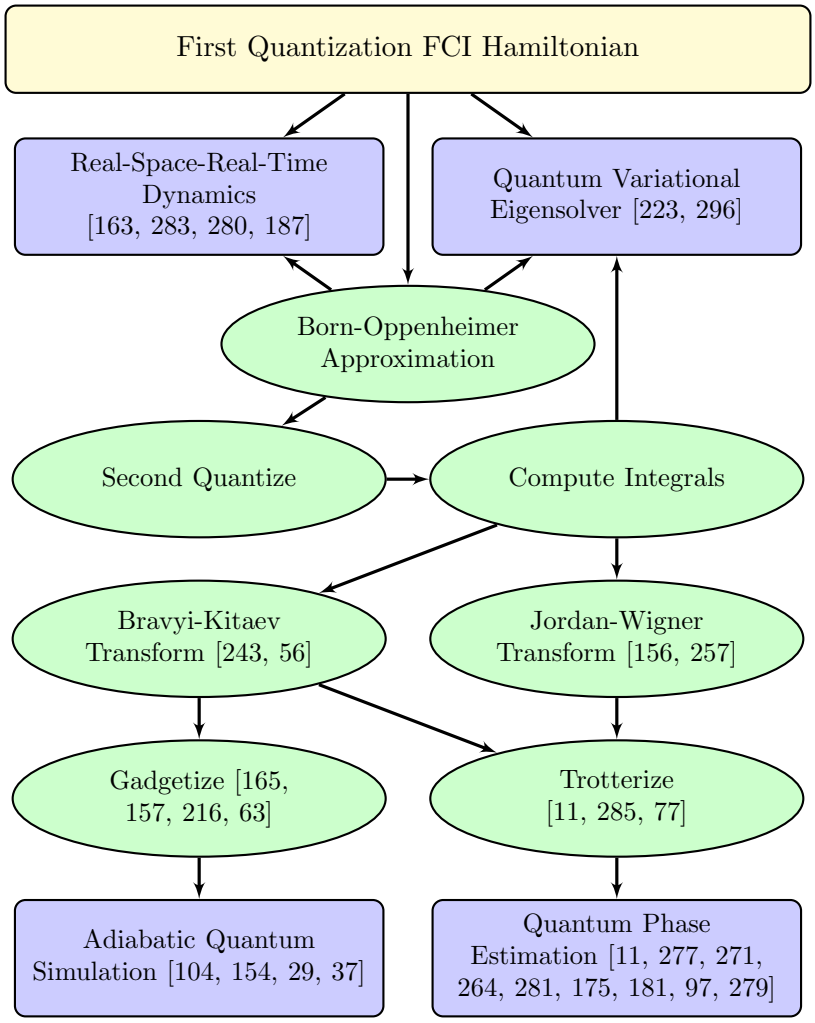



Figure 7.1: A diagram relating several different approaches to the quantum simulation of quantum chemistry with the procedures and approximations implicit in each approach. Some of these approaches have been demonstrated experimentally using quantum information processors.

The worst case complexity of generic model chemistries (e.g. local fermionic problems studied with density functional theory) has been shown to be in the quantum mechanical equivalent of NP-Complete, QMA-Complete [242, 286]. However, the subset of these generic models which correspond to stable molecules, or to unstable configurations of chemical interest such as transition states, is small and structured. Just as with adiabatic optimization, it does not matter if molecular electronic structure is QMA-Complete so long as the average instance can be solved (or even approximated) efficiently. In this case we also have considerable heuristic evidence that molecules are able to find their ground state configurations rapidly: these are the configurations in which they naturally occur. Similarly, unstable transition states of interest occur in natural processes. Given that simulation of time evolution on a quantum computer is efficient, we conjecture that simulation of the natural processes that give rise to these states will also be practical.

The proofs that Local Hamiltonian (a decision problem capturing the complexity of finding the ground state energy) is QMA-Complete relies on the construction of various specific Hamiltonians that can represent any possible instance of any problem in QMA. In general, these Hamiltonians possess couplings between more than two qubits. Hamiltonians which contain many-body interactions of order $k$ and lower are referred to as $k$-local Hamiltonians; experimentally programmable couplings are 2-local. The original formulation by Kitaev was $(\log n)$-local, he then reduced this to 5-local and that result was subsequently reduced to 3-local. To reduce 3-local Hamiltonians to 2-local Hamiltonians "perturbative gadgets" were introduced by Kempe *et al.* [165], which can embed a $k$-local Hamiltonian in a subspace of a 2-local Hamiltonian using ancilla qubits. In the past decade, a growing body of work has pushed the development of different gadgets which embed various target Hamiltonians with various tradeoffs in the resources required [216, 157, 54, 38, 98, 16, 63].

Embedding problems in realizable Hamiltonians requires careful consideration of the availability of experimental resources. One consideration is that many-body qubit interactions cannot be directly realized experimentally. Another factor is the "control precision" of the Hamiltonian which is the dynamic range of field values which a device must be able to resolve in order to embed the intended eigenspectrum to a desired accuracy. This resource is especially important for molecular electronic structure Hamiltonians as chemists are typically interested in acquiring chemical accuracy (0.04 eV). Control precision is often the limiting factor when a Hamiltonian contains terms with coefficients which vary by several orders of magnitude. Other considerations include the number of qubits available as well as the connectivity and type of qubit couplings.

In this paper, we describe a scalable method which allows for the application of the quantum adiabatic algorithm to a programmable physical system encoding the molecular electronic Hamiltonian. Our method begins with the second quantized representation of molecular electronic structure in which the Hamiltonian is represented with fermionic creation and annihilation operators. The first step in our protocol is to convert the fermionic Hamiltonian to a qubit Hamiltonian using the Bravyi-Kitaev transformation [56, 243]. We show that using the Bravyi-Kitaev transformation instead of the Jordan-Wigner transformation is necessary for avoiding exponential control precision requirements in an experimental setting. Next, we show a new formulation of perturbative gadgets motivated by [165, 63] that allows us to remove all terms involving $YY$ couplings in a single gadget application (note that throughout this paper we use $X$, $Y$ and $Z$ to denote the Pauli matrices and these operators are defined to act as identity on unlabeled registers so that the dot product $Y_iY_j$ is understood to represent the tensor product $Y_i \otimes Y_j$). Finally, we apply the gadgets described in [157] to produce a 2-local Hamiltonian with only $ZZ$, $XX$ and $ZX$ couplings.

The paper is organized as follows. In the first section we review the second quantized formulation of the electronic structure problem. Next we give the mapping of this problem to qubits. In the third section we introduce the gadgets that we will use for locality reduction. Finally, we apply our procedure to a simple example: molecular hydrogen in a minimal basis. We close the paper with some discussion and directions for future work.

## 7.2 Second Quantization

We begin by writing down the full configuration interaction (FCI) Hamiltonian in the occupation number basis. We define spin orbitals as the product of a spin function (representing either spin up or spin down) and a single-electron spatial function (usually molecular orbitals produced from a Hartree-Fock calculation). For example, in the case of molecular hydrogen there are two electrons and thus, two single-electron molecular orbitals, $|\psi_1\rangle$ and $|\psi_2\rangle$. Electrons have two possible spin states, $|\alpha\rangle$ (spin up) and $|\beta\rangle$ (spin down). The four spin orbitals for molecular hydrogen are therefore, $|\chi_0\rangle = |\psi_1\rangle|\alpha\rangle$, $|\chi_1\rangle = |\psi_1\rangle|\beta\rangle$, $|\chi_2\rangle = |\psi_2\rangle|\alpha\rangle$, and $|\chi_3\rangle = |\psi_2\rangle|\beta\rangle$.

The occupation number basis is formed from all possible configurations of $n$ spin orbitals which are each either empty or occupied. We represent these vectors as a tensor product of individual spin orbitals written as $|f_{n-1}...f_0\rangle$ where $f_j \in \mathbb{B}$ indicates the occupation of spin orbital $|\chi_j\rangle$. Any interaction between electrons can be represented as some combination of creation and annihilation operators $a_j^\dagger$ and $a_j$ for $\{j \in \mathbb{Z} | 0 \leq j < n\}$. Because fermionic wavefunctions must be antisymmetric with respect to particle label exchange, these operators must obey the fermionic anti-commutation relations,

$$[a_j, a_k]_+ = [a_j^\dagger, a_k^\dagger]_+ = 0, \quad [a_j, a_k^\dagger]_+ = \delta_{jk}\mathbf{1}. \tag{7.1}$$

With these definitions we write the second-quantized molecular electronic Hamiltonian,

$$H = \sum_{i,j} h_{ij} a_i^\dagger a_j + \frac{1}{2} \sum_{i,j,k,l} h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l. \tag{7.2}$$

The coefficients $h_{ij}$ and $h_{ijkl}$ are single and double electron overlap integrals which are precomputed classically. The number of distinct integrals scale as $O\left(n^4\right)$ in the number of molecular orbitals $n$.

## 7.3 Qubit Representation

The next step in our reduction will be to represent our fermionic wavefunction in terms of qubits. We use the direct mapping introduced in [11] that maps an occupancy state to a qubit basis state. Using Pauli operators we can represent qubit raising and lowering operators as,

$$\begin{aligned} Q_j^+ &= |1\rangle\langle 0| = \frac{1}{2}\left(X_j - iY_j\right), \\ Q_j^- &= |0\rangle\langle 1| = \frac{1}{2}\left(X_j + iY_j\right). \end{aligned} \tag{7.3}$$

However, these operators do not obey the fermionic commutation relations given in Eq. 7.1. To write qubit operators that obey the commutation relations in Eq. 7.1, we could use the Jordan-Wigner transformation [156, 257, 11].

Unfortunately, the Jordan-Wigner transformation is not a scalable way to reduce electronic structure to an experimentally realizable Hamiltonian for AQC. This is because the Jordan-Wigner transformation introduces $k$-local interaction terms into the Hamiltonian and $k$ grows linearly in the system size. Prima facie, this is not a major problem because there exist theoretical tools known as perturbative gadgets which allow for reductions in interaction order. However, in all known formulations of perturbative gadgets, control precision increases exponentially in $k$. Thus, the linear locality overhead introduced by the

Jordan-Wigner transformation translates into an exponential control precision requirement in the reduction.

An alternative mapping between the occupation number basis and qubit representation, known as the Bravyi-Kitaev transformation, introduces logarithmic locality overhead [56, 243]. Two pieces of information are required in order to correctly construct creation and annihilation operators that act on qubits and obey the fermionic commutation relations. First, the occupancy of each orbital must be stored. Second, parity information must be stored so that for a pair of orbitals, it is possible to determine the parity of the occupancy of the orbitals that lie between them. This parity determines the phase which results from exchanging the occupancy of the two orbitals.

The occupation number basis stores the occupation directly in the qubit state (hence the name). This implies that occupancy is a fully local variable in this basis; one may determine the occupancy of an orbital by measuring a single qubit. However, this also implies that the parity information is completely non-local. It is this fact that determines the structure of the qubit creation and annihilation operators in the Jordan-Wigner transformation. Each such operator changes the state of a single qubit $j$ (updating the occupancy information) but also acts on all qubits with indices less than $j$ to determine the parity of their occupancy. This results in qubit operators, expressed as tensor products of Pauli matrices, that contain strings of $Z$ operators whose length grows with the number of qubits. One could consider storing the parity information locally, so that the qubit basis states store sums of orbital occupancies. Then determination of parity requires a single qubit operation. However, updating occupancy information requires updating the state of a number of qubits that again grows with the number of qubits. Hence this "parity basis" construction offers no advantage over the Jordan Wigner transformation [56].

The Bravyi-Kitaev transformation offers a middle ground in which both parity and occupancy information are stored non-locally, so neither can be determined by measurement of a single qubit [56, 243]. Both parity and occupancy information can be accessed by acting on a number of qubits that scales as the logarithm of the number of qubits. This logarithmic scaling makes the proposed mapping of electronic structure to a 2-local qubit Hamiltonian efficient.

The consequences of this mapping, originally defined in [56], were computed for electronic structure in [243]. That work defines several subsets of qubits in which the parity and occupancy information is stored. The occupancy information is stored in the update set, whereas the parity information is stored in the parity set. These sets are distinct and their size is strictly bounded above by the logarithm base two of the number of qubits. The total number of qubits on which a qubit creation and annihilation operator may act can be a multiple of the logarithm base two of the number of qubits. However, this multiple is irrelevant from the point of view of the scalability of the construction. Using the Bravyi-Kitaev transformation, the spin Hamiltonian for molecular hydrogen in the minimal (STO-3G) basis, as reported in [243], is given by

$$H_{\mathrm{H}_2} = f_0\mathbf{1} + f_1 Z_0 + f_2 Z_1 + f_3 Z_2 + f_1 Z_0 Z_1 + f_4 Z_0 Z_2 + f_5 Z_1 Z_3 + f_6 X_0 Z_1 X_2 + f_6 Y_0 Z_1 Y_2$$
$$+ f_7 Z_0 Z_1 Z_2 + f_4 Z_0 Z_2 Z_3 + f_3 Z_1 Z_2 Z_3 + f_6 X_0 Z_1 X_2 Z_3 + f_6 Y_0 Z_1 Y_2 Z_3 + f_7 Z_0 Z_1 Z_2 Z_3 \quad (7.4)$$

where the integral values (in Hartree) are,

$$f_0 = -0.81261,\ f_1 = 0.17120,\ f_2 = 0.16862,\ f_3 = -0.22278, \quad (7.5)$$
$$f_4 = 0.12055, f_5 = 0.17435,\ f_6 = 0.04532,\ f_7 = 0.16587.$$

In general, the Bravyi-Kitaev transformation applied to electronic structure produces an $n$-qubit Hamiltonian which is $(\log n)$-local, and has $n^4$ real terms. This implies that each

term has an even number of $Y$ terms, or none.

## 7.4 Hamiltonian Gadgets

In order to embed electronic structure in an experimentally realizable Hamiltonian, we define a scalable methodology for transforming our $(\log n)$-local qubit Hamiltonian into a 2-local Hamiltonian with only $ZZ$, $XX$ and $XZ$ interaction terms. In this section we will describe tools known as "gadgets" which allow us to simulate the target Hamiltonian with these interactions.

Hamiltonian gadgets provide a method for embedding the eigenspectra (and sometimes eigenvectors) of an $n$-qubit "target" Hamiltonian, denoted by $H_{\text{target}}$, in a restricted (typically low-energy) subspace of a more constrained $(N > n)$-qubit "gadget" Hamiltonian, denoted by $\widetilde{H}$. To illustrate the general idea of gadgets, we describe how a 2-local Hamiltonian can embed a $k$-local Hamiltonian. Suppose that we have a gadget Hamiltonian, $\widetilde{H}$, which contains only 2-local terms which act on $N = n + a$ qubits. Then,

$$\widetilde{H} = \sum_{i=1} f_i O_i, \quad \widetilde{H}|\psi_i\rangle = \widetilde{\lambda}_i|\widetilde{\psi}_i\rangle, \tag{7.6}$$

where $\{f_i\}$ are scalar coefficients, $\widetilde{\lambda}_j$ and $|\widetilde{\psi}_i\rangle$ are the eigenvectors and eigenvalues of $\widetilde{H}$, and $\{O_i\}$ are the 2-local interaction terms of the physical Hamiltonian. We choose our interaction terms to be Hilbert-Schmidt orthogonal so that $\text{Tr}\left[O_i O_j\right] = 2^n \delta_{i,j}$. We now define an *effective Hamiltonian* which has support on the lowest $2^n$ states of the gadget,

$$H_{\text{eff}} \equiv \sum_{i=0}^{2^n-1} \widetilde{\lambda}_i |\widetilde{\psi}_i\rangle\langle\widetilde{\psi}_i| = \sum_{i=1} f_i O_i \otimes \Pi. \tag{7.7}$$

Here $\Pi$ is a projector onto a particular state (usually the lowest energy state) of the $a$ ancilla qubits and the $\{O_i\}$ are a Hilbert-Schmidt orthogonal operator basis for operators on the space of the $n$ logical qubits. In other words, the most general representation of $H_{\text{eff}}$ is an

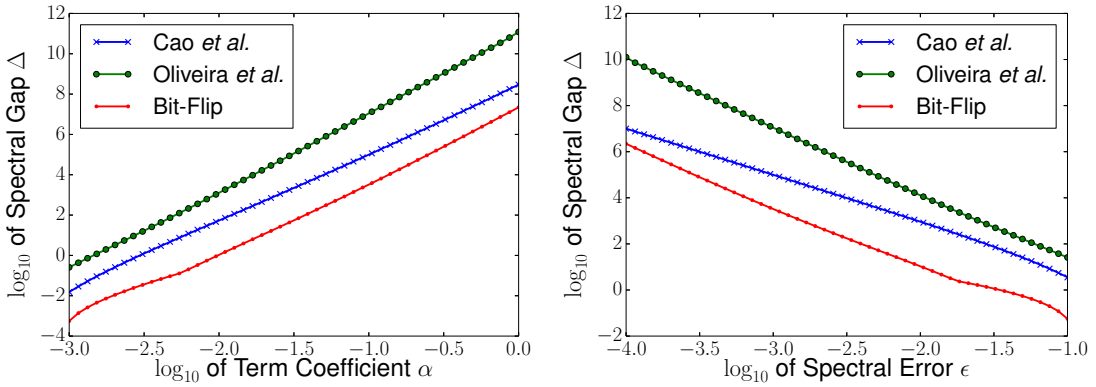



Figure 7.2: Numerics comparing the minimum spectral gaps required to reduce the term $\alpha X_1 Y_2 Z_3$ to 2-local with an error in the eigenspectrum of at most $\epsilon$. On the left, $\epsilon$ is fixed at 0.001 and gaps are plotted as a function of $\alpha$. On the right, $\alpha$ is fixed at 0.1 and gaps are plotted as a function of $\epsilon$. Here we compare the bit-flip construction [165, 157], the Oliveira and Terhal construction [216] and an improved variant on Oliveira and Terhal by Cao *et al.* [63].

expansion of all possible tensor products acting on the logical qubits. In general, there is no reason why $f_i = 0$ on all non-2-local terms. Therefore a 2-local gadget on $N = n + a$ qubits can embed a $(k > 2)$-local, $n$-qubit Hamiltonian using $a$ ancilla bits.

The use of perturbation theory to derive Hamiltonian gadgets was introduced by Kempe *et al.* in their canonical proof showing that 2-Local Hamiltonian is QMA-Complete [165]. Their construction, which we refer to as the "bit-flip construction" for reasons that will become obvious later on, was analyzed by Jordan and Farhi using a formulation of perturbation theory due to Bloch [157]. Other perturbative gadget constructions were introduced by Oliveira and Terhal to prove the QMA-Completeness of Hamiltonian on a square lattice [216]. Following this work, Biamonte and Love used gadgets to show that $XX$ and $ZZ$, or $XZ$ couplings alone, suffice for the QMA-Completeness of 2-local Hamiltonian [38]. Several other papers improve these gadgets from an experimental perspective and introduce novel constructions which are compatible with the protocol developed here [54, 98, 16, 63]. We

note that different types of gadgets may have specific advantages when designing Hamiltonians for specific hardware. Results from [63] suggest that there is a rough tradeoff between the number of ancillae required and the amount of control precision required. For instance, Figure 7.2 indicates that bit-flip gadgets require less control precision than other gadget constructions (but generally more ancillae). In this paper we focus on the bit-flip family of gadgets.

Although we employ the perturbation theory approach here, it does require a high degree of control precision and should be avoided when possible. We point out that when the Hamiltonian is entirely diagonal there are exact gadgets [16] which can embed the ground state with far less control precision and often far fewer ancillae but in a way that does not necessarily conserve the gap scaling. Moreover, "frustration-free" gadgets have been used extensively in proofs of the QMA-Completeness of various forms of quantum satisfiability, and in restricting the necessary Hamiltonian terms for universal adiabatic quantum computing [204, 203, 123, 70].

While several types of perturbation theory have been used to derive these gadgets, we closely follow the approach and notation of Kempe *et al.* [165]. We wish to analyze the spectrum of the gadget Hamiltonian, $\widetilde{H} = H + V$ for the case that the norm of the perturbation Hamiltonian, $V$, is small compared to the spectral gap between the ground state and first excited state of the unperturbed Hamiltonian, $H$. To accomplish this we use the Green's function of $\widetilde{H}$,

$$\widetilde{G}(z) \equiv \left(z\mathbf{1} - \widetilde{H}\right)^{-1} = \sum_j \frac{|\widetilde{\psi}_j\rangle\langle\widetilde{\psi}_j|}{z - \widetilde{\lambda}_j}. \tag{7.8}$$

We also define $G(z)$ using the same expression except with $H$ instead of $\widetilde{H}$. Further, let $\mathcal{H} = \mathcal{L}_+ \oplus \mathcal{L}_-$ be the Hilbert space of $\widetilde{H}$ where $\mathcal{L}_+$ is the "high-energy" subspace spanned by eigenvectors of $\widetilde{H}$ with eigenvalues $\widetilde{\lambda} \geq \lambda_*$ and $\mathcal{L}_-$ is the complementary "low-energy"

subspace, spanned by eigenvectors of $\widetilde{H}$ corresponding to eigenvalues of $\widetilde{\lambda} < \lambda_*$. Let $\Pi_\pm$ correspond to projectors onto the support of $\mathcal{L}_\pm$. In a representation of $\mathcal{H} = \mathcal{L}_+ \oplus \mathcal{L}_-$, all the aforementioned operators $V$, $H$, $\widetilde{H}$, $G(z), \widetilde{G}(z)$ are block-diagonal so we employ the notation that $A_{\pm\pm} = \Pi_\pm A \Pi_\pm$ and,

$$A = \begin{pmatrix} A_+ & A_{+-} \\ A_{-+} & A_- \end{pmatrix}. \tag{7.9}$$

Finally, we define the operator function known as the *self-energy*,

$$\Sigma_-(z) \equiv z\mathbf{1}_- - \widetilde{G}_-^{-1}(z). \tag{7.10}$$

We use this notation to restate the "gadget theorem".

**Theorem 7.4.1.** *Theorem 6.2 in [165]. Assume that $H$ has a spectral gap $\Delta$ around the cutoff $\lambda_*$; i.e. all of its eigenvalues are in $(-\infty, \lambda_-] \cup [\lambda_+, +\infty)$ where $\lambda_+ = \lambda_* + \Delta/2$ and $\lambda_- = \lambda_* - \Delta/2$. Assume that $\|V\| \leq \Delta/2$. Let $\epsilon > 0$ be arbitrary. Assume there exists an operator $H_{eff}$ such that $\lambda(H_{eff}) \subset [c, d]$ for some $c < d < \lambda_* - \epsilon$ and, moreover, the inequality $\|\Sigma_-(z) - H_{eff}\| \leq \epsilon$ holds for all $z \in [c - \epsilon, d + \epsilon]$. Then each eigenvalue $\widetilde{\lambda}_j$ of $\widetilde{H}_-$ is $\epsilon$-close to the $j^{th}$ eigenvalue of $H_{eff}$.*

Theorem 1 assures us that the eigenspectrum of the self-energy provides an arbitrarily good approximation to the eigenspectrum of the low-energy subspace of the gadget Hamiltonian. This is useful because the self-energy admits a series expansion,

$$\Sigma_-(z) = H_- + V_- + \sum_{k=2}^{\infty} V_{-+} G_+ (V_+ G_+)^{k-2} V_{+-}. \tag{7.11}$$

Using $G_+ = (z - \Delta)^{-1} \mathbf{1}_+$ and $H_- = 0$, we focus on the range $z = O(1) \ll \Delta$ and find that,

$$H_{\text{eff}} \approx V_- + \frac{1}{\Delta} \sum_{k=2}^{\infty} V_{-+} \left(\frac{V_+}{\Delta}\right)^{k-2} V_{+-}. \tag{7.12}$$

We use this effective Hamiltonian to approximate our $k$-local target Hamiltonian, which we now specify. The terms in our target Hamiltonian will have a locality that scales logarithmically with the number of orbitals. We may write such a term:

$$T = \bigotimes_{i=0}^{k-1} O_i \ : \ O_i \in \{X_i, Y_i, Z_i\} \quad \forall \ i. \tag{7.13}$$

One can always apply gadgets term by term to reduce locality; however, this may not be the optimal procedure. In addition, we are interested in replacing even tensor powers of the $Y$ operator. For both these reasons we consider a slightly more general form of term as a target for gadgetization. We use the fact that it is only the commuting nature of the $\{O_i\}$ that is important for the gadget to function. We therefore write our target term as a product of $k$ commuting operators, which includes the special case in which it is a product of $k$ operators acting on distinct tensor factors,

$$T' = \prod_{i=0}^{k-1} O_i \ : \ [O_i, O_j] = 0 \quad \forall \ \{i, j\} \tag{7.14}$$

Hence, we can represent the target Hamiltonian as a sum of $r$ terms which are the product of $k$ commuting operators,

$$H_{\text{target}} = H_{\text{else}} + \sum_{s=1}^{r} \prod_{i=0}^{k-1} O_{s,i} \tag{7.15}$$

where all $\{O_{s,i}\}$ commute for a given $s$ and $H_{\text{else}}$ can be realized directly by the physical Hamiltonian. While previous formulations of bit-flip gadgets [165, 157, 63] have gadgetized operators acting on distinct tensor factors, it is only necessary that the operators commute. Their action on distinct tensor factors is sufficient but not necessary for the gadget construction. We take advantage of this property in order to realize $YY$ terms without access to such couplings by making the substitution, $Y_iY_j \rightarrow -X_iX_jZ_iZ_j$. Since $X_iX_j$ commutes with $Z_iZ_j$, we can create this effective interaction with a bit-flip gadget. For instance, suppose we have the term, $Z_0Y_1Y_2$. We gadgetize the term $A \cdot B \cdot C$ where $A = Z_0$,

$B = -X_1X_2$, and $C = Z_1Z_2$ and all operators $A, B, C$ commute. We note that another approach to removing $YY$ terms is explained in [63].

We now introduce the form of the penalty Hamiltonian that acts only on the ancilla qubits. Bit-flip gadgets introduce an ancilla system which has two degenerate ground-states, usually taken to be $|111...\rangle_u$ and $|000...\rangle_u$ where $u$ indicates that these kets refer to an ancilla space. For each of the $r$ terms we use a separate ancilla system of the form,

$$H_s = \frac{\Delta}{2\,(k-1)} \sum_{0 \le i < j \le k-1} \left(\mathbf{1} - Z_{u_{s,i}} Z_{u_{s,j}}\right). \tag{7.16}$$

Again, we use $u$ to indicate that operators act on an ancilla; e.g. the label $u_{3,2}$ indicates the ancilla corresponding to $O_{3,2}$ (the second operator in the third term). For each term we follow Farhi and Jordan in introducing an ancilla system connected by a complete graph with equal and negative edge weights. Thus, the ground state of the ancilla system is spanned by $|111...\rangle_u$ and $|000...\rangle_u$.

Next, we introduce the perturbation Hamiltonian,

$$V = H_{\text{else}} + \Lambda + \mu \sum_{s=1}^{r} \sum_{i=0}^{k-1} O_{s,i} X_{u_{s,i}}, \tag{7.17}$$

where $\mu = \sqrt[k]{\frac{\Delta^{k-1}}{k!}}$ and $\Lambda$ is a 2-local operator on logical bits which will be discussed later. The effect of this Hamiltonian on the low energy subspace is to introduce virtual excitations into the high energy space that modify the low energy effective Hamiltonian. Only terms which start and end in the ground state contribute to the perturbation series for the self-energy (see, for example, Figure 7.3). Thus, the gadget will produce the target term at order $k$ in which a transition between the two degenerate ground states of the ancillae requires that each of the $X_u$ terms in the perturbation act exactly once to flip all $r \cdot k$ bits from one ground state to the other. Crucially, the order in which the ancillae are flipped does not

matter since the operators $O_{s,i}$ commute for a given $s$. The complete gadget is

$$\widetilde{H} = \Lambda + H_{\text{else}} + \sum_{s=1}^{r} \left[ \mu \sum_{i=0}^{k-1} O_{s,i} X_{u_{s,i}} + \frac{\Delta}{2\,(k-1)} \sum_{0 \le i < j \le k-1} \left( \mathbf{1} - Z_{u_{s,i}} Z_{u_{s,j}} \right) \right]. \tag{7.18}$$

and is related to the target Hamiltonian and effective Hamiltonian by,

$$\widetilde{H}_{-} = H_{\text{target}} \otimes \Pi_{-} = H_{\text{eff}} \tag{7.19}$$

for the appropriate choice of $\Lambda$ and $\Delta \gg \|V\|$ where $\Pi_{-}$ projects onto the ancillae ground space,

$$\Pi_{-} = |000\rangle \langle 000|_u + |111\rangle \langle 111|_u \,. \tag{7.20}$$

To illustrate the application of such a gadget and demonstrate how $\Lambda$ is chosen, we scalably reduce the locality of molecular hydrogen and remove all $Y$ terms in the next section.

For the example $H_{\text{target}} = A \cdot B \cdot C + H_{\text{else}}$, the perturbation is

$$V = \mu A X_a + \mu B X_b + \mu C X_c + H_{\text{else}} + \Lambda. \tag{7.21}$$

Its components in the low energy subspace, as in the block diagonal representation of Eq. 7.9 is:

$$V_{-} = (H_{\text{else}} + \Lambda) \otimes (|000\rangle\langle 000|_u + |111\rangle\langle 111|_u) \,. \tag{7.22}$$

The projection into the high energy subspace is:

$$\begin{aligned} V_{+} = {} & (H_{\text{else}} + \Lambda) \otimes \left( \sum_{\{a,b,c\} \in \mathbb{B}^3} |a,b,c\rangle\langle a,b,c|_u \right) - V_{-} \\ & + \mu A \otimes (|0,1,0\rangle\langle 1,1,0|_u + |1,1,0\rangle\langle |0,1,0|_u + |0,0,1\rangle\langle 1,0,1|_u + |1,0,1\rangle\langle 0,0,1|_u) \\ & + \mu B \otimes (|1,0,0\rangle\langle 1,1,0|_u + |1,1,0\rangle\langle |1,0,0|_u + |0,0,1\rangle\langle 0,1,1|_u + |0,1,1\rangle\langle 0,0,1|_u) \\ & + \mu C \otimes (|1,0,0\rangle\langle 1,0,1|_u + |1,0,1\rangle\langle |1,0,0|_u + |0,1,0\rangle\langle 0,1,1|_u + |0,1,1\rangle\langle 0,1,0|_u) \,. \end{aligned} \tag{7.23}$$

The projections coupling the low and high energy subspaces are:

$$V_{+-} = \mu A \otimes (|1,0,0\rangle\langle 0,0,0|_u + |0,1,1\rangle\langle 1,1,1|_u) \tag{7.24}$$
$$+ \mu B \otimes (|0,1,0\rangle\langle 0,0,0|_u + |1,0,1\rangle\langle 1,1,1|_u) + \mu C \otimes (|0,0,1\rangle\langle 0,0,0|_u + |1,1,0\rangle\langle 1,1,1|_u)$$

and $V_{-+} = (V_{+-})^\dagger$. Substituting these values into Eq. 7.12 we see that at order $k = 3$ a term appears with the following form,

$$\frac{1}{\Delta^2} V_{-+} V_+ V_{+-} = \frac{\mu^3}{\Delta^2} (ABC + ACB + BCA + CAB + BAC + CBA) \to ABC. \tag{7.25}$$

These terms arise because all ancilla qubits must be flipped and there are six ways of doing so, representing 3! (in general this will be $k!$ for a gadget with $k$ ancillae) combinations of the operators. These six terms are represented diagrammatically in Figure 7.3. Note that it is the occurrence of all orderings of the operators $A$, $B$ and $C$ that imposes the requirement that these operators commute. Hence, in order to realize our desired term we see that $\mu = \sqrt[k]{\frac{\Delta^{k-1}}{k!}}$. A few competing processes occur which contribute unwanted terms but these terms either vanish with increasing spectral gap $\Delta$, or they can be removed exactly by introducing terms into the compensation term $\Lambda$. A simple way to compute $\Lambda$ is to evaluate the perturbation series to order $k$ and choose $\Lambda$ so that problematic terms disappear.

At higher orders we encounter "cross-gadget contamination" which means that processes occur involving multiple ancilla systems, causing operators from different terms to interact. For a 3-operator gadget, such terms will always only contribute at order $O\left(\Delta^{-3}\right)$. In reductions which require going to higher orders, these terms do not necessarily depend on $\Delta$, and so may introduce unwanted terms into the effective Hamiltonian. For instance, Figure 7.4 shows an example of the four processes which occur at fourth order for a multiple term, 4-operator reduction. The diagrams involving multiple ancilla registers are examples of cross-gadget contamination.

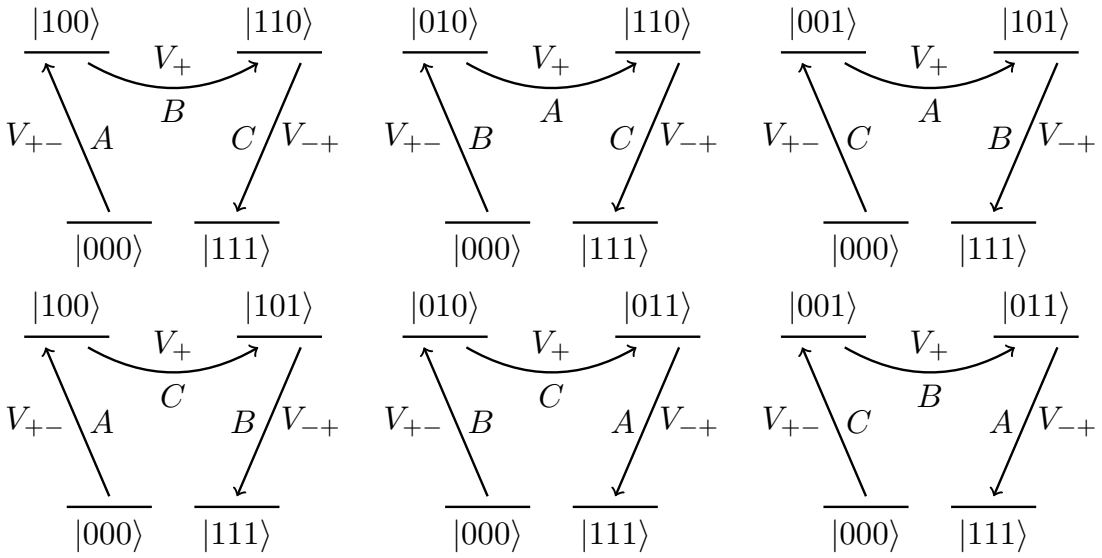



Figure 7.3: The six equivalent bit-flip processes at third order which produce the effective interaction $A \cdot B \cdot C$. Each of these diagrams also occurs backwards on the part of the ground state in $|111\rangle$.

However, if terms are factored into tensor products of operators that square to the identity (as is the case for products of Pauli operators, which is always possible), cross-gadget contamination can only contribute a constant shift to the energy which can be compensated for in $\Lambda$. This is because any process contributing to the perturbation series which does not transition between the two different ground states must contain an even multiple of each operator and if we choose to act on the non-ancilla qubits with operators that square to identity we obtain only a constant shift. Consider the two cross-gadget terms represented in these diagrams: $A_1 C_2^2 A_1 = A_1 \mathbf{1} A_1 = \mathbf{1}$ and $D_2 B_1 D_2 B_1 = (D_2 B_1)^2 = \mathbf{1}$. At even higher orders, *individual* cross-gadget terms might not equal a constant shift (i.e. the sixth order term $A_1 A_2 A_3 A_2 A_1 A_3$) but the occurrence of all combinations of operators and the fact that all Pauli terms either commute or anti-commute will guarantee that such terms disappear. In the sixth order example, if $[A_1, A_2] = 0$ then $A_1 A_2 A_3 A_2 A_1 A_3 = A_1 A_2 A_3 A_1 A_2 A_3 = (A_1 A_2 A_3)^2 = \mathbf{1}$, otherwise $[A_1, A_2]_+ = 0$ which

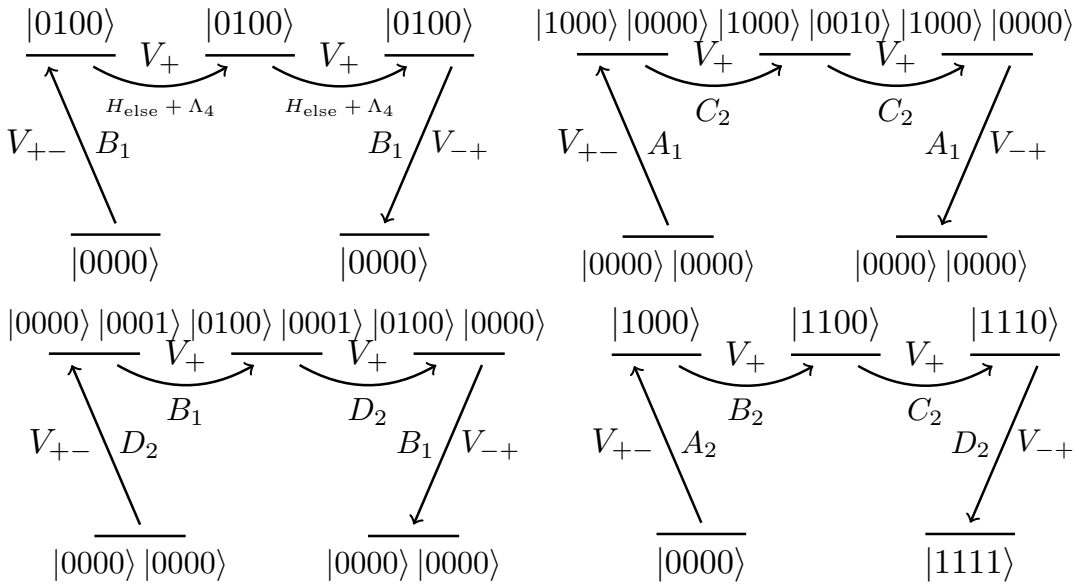



Figure 7.4: Diagrams showing an example of each of the four processes at fourth order. In the upper left is the process $B_1 \left(H_{\text{else}} + \Lambda\right)^2 B_1$. In the upper right is the process $A_1 C_2^2 A_1$. In the lower left is the process $D_2 B_1 D_2 B_1$. In the lower right is the process $A_2 B_2 C_2 D_2$.

implies that $A_1 A_2 A_3 A_2 A_1 A_3 + A_1 A_2 A_3 A_1 A_2 A_3 = 0$.

## 7.5 Example Problem: Molecular Hydrogen

We begin by factoring and rewriting the $k$-local molecular hydrogen Hamiltonian from Eq. 7.4 into a 4-local part and a 2-local part so that $H_{\mathrm{H}_2} = H_{4\mathrm{L}} + H_{2\mathrm{L}}$ where,

$$H_{4\mathrm{L}} = (f_4 Z_0 + f_3 Z_1)\, Z_2 Z_3 + (Z_1 + Z_1 Z_3)\,(f_6 X_0 X_2 + f_6 Y_0 Y_2 + f_7 Z_0 Z_2) \tag{7.26}$$

$$H_{2\mathrm{L}} = f_0 \mathbf{1} + f_2 Z_1 + f_3 Z_2 + f_4 Z_0 Z_2 + f_5 Z_1 Z_3 + f_1 Z_0\,(\mathbf{1} + Z_1)\,. \tag{7.27}$$

In order to reduce $H_{\mathrm{H}_2}$ to a 2-local $ZZ/XX/XZ$-Hamiltonian we further factor $H_{4\mathrm{L}}$ to remove $YY$ terms,

$$\begin{aligned} H_{4\mathrm{L}} &= \underbrace{(f_4 Z_0 + f_3 Z_1)}_{A_1}\underbrace{Z_2}_{B_1}\underbrace{Z_3}_{C_1} + \underbrace{f_7 Z_0}_{A_2}\underbrace{Z_2}_{B_2}\underbrace{(Z_1 + Z_1 Z_3)}_{C_2} \\ &\quad + \underbrace{f_6 X_0 X_2}_{A_3}\underbrace{(\mathbf{1} - Z_0 Z_2)}_{B_3}\underbrace{(Z_1 + Z_1 Z_3)}_{C_3} \\ &= A_1 B_1 C_1 + A_2 B_2 C_2 + A_3 B_3 C_3. \end{aligned} \tag{7.28}$$

Within each term, the operators all commute so that $[A_i, B_i] = [A_i, C_i] = [B_i, C_i] = 0$. We emphasize that factoring terms into commuting operators is always possible and necessary in order for bit-flip gadgets to work correctly.

Each of the operators defined in Eq. 7.28 will have a corresponding ancilla qubit labelled to indicate the operator with which it is associated, e.g. the ancilla for operator $B_2$ has label $b_2$. Our unperturbed Hamiltonian is a sum of fully connected ancilla systems in which each ancilla system corresponds to a term,

$$\begin{aligned} H_1 = \frac{9\Delta_1}{4}\mathbf{1} - \frac{\Delta_1}{4}\,(&Z_{a_1} Z_{b_1} + Z_{a_1} Z_{c_1} + Z_{b_1} Z_{c_1} \\ &+ Z_{a_2} Z_{b_2} + Z_{a_2} Z_{c_2} + Z_{b_2} Z_{c_2} + Z_{a_3} Z_{b_3} + Z_{a_3} Z_{c_3} + Z_{b_3} Z_{c_3})\,. \end{aligned} \tag{7.29}$$

The spectral gap and Hamiltonian have the subscript "1" to associate them with the first

of two applications of perturbation theory. We perturb the ancilla system with the Hamiltonian,

$$V_1 = \mu_1 \left( A_1 X_{a_1} + B_1 X_{b_1} + C_1 X_{c_1} + A_2 X_{a_2} + B_2 X_{b_2} \right. \tag{7.30}$$
$$\left. + C_2 X_{c_2} + A_3 X_{a_3} + B_3 X_{b_3} + C_3 X_{c_3} \right) + H_{2\mathrm{L}} + \Lambda_1$$

where $\mu_1 = \sqrt[3]{\frac{\Delta_1^2}{6}}$ and $\Lambda_1$ is a 2-local compensation Hamiltonian acting on the logical qubits only. Later on, $\Lambda_1$ will be chosen to cancel extraneous terms from the perturbative expansion. The interaction terms involving $A$, $B$, and $C$ will arise at third order ($V_{-+}V_{+}V_{+-}$) from processes which involve a transition between the two degenerate ground states of the ancilla systems. This occurs at third order because to make the transition $|000\rangle \rightleftharpoons |111\rangle$, we must flip all three ancilla bits in each term by applying the operators $X_a$, $X_b$, and $X_c$. Since these operators are coupled to $A$, $B$, and $C$, sequential action of bit flip operators yields our desired term. Because the operators commute, the order of the bit flipping does not matter. We now calculate the effective Hamiltonian using the perturbative expansion of the self-energy from Eq. 7.12.

### 7.5.1 Second Order

The only processes which start in the ground state and return to the ground state at second order are those which flip a single bit and then flip the same bit back. Thus, effective interactions are created between each operator and itself,

$$-\frac{1}{\Delta_1} V_{-+} V_{+-} = -\frac{\mu_1^2}{\Delta_1} \left( A_1^2 + B_1^2 + C_1^2 + A_2^2 + B_2^2 + C_2^2 + A_3^2 + B_3^2 + C_3^2 \right) \tag{7.31}$$
$$= -\sqrt[3]{\frac{\Delta_1}{36}} \left[ \left( 9 + f_3^2 + f_4^2 + f_6^2 + f_7^2 \right) \mathbf{1} + 2 f_3 f_4 Z_0 Z_1 - 2 Z_0 Z_2 + 4 Z_3 \right].$$

These processes are shown in Figure 7.5.

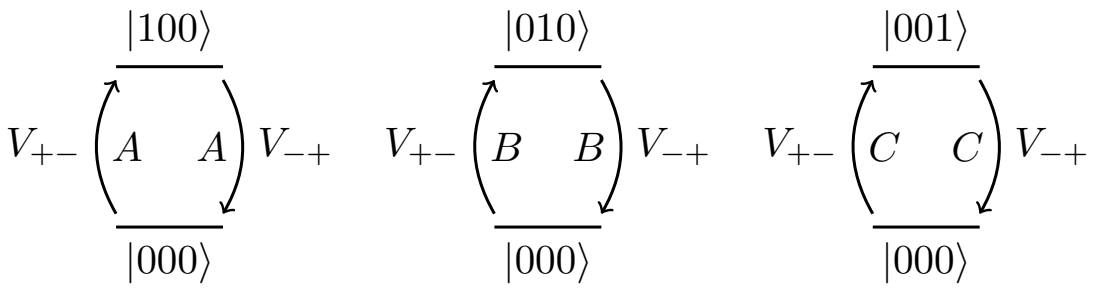



Figure 7.5: The three bit-flip processes at second order. These occur for each term. Note that each of these diagrams occurs in reverse for the part of the ground state in $|111\rangle$.

The second order effective Hamiltonian at large $\Delta_1$ is,

$$H_{\text{eff}}^{(2)} = H_{2\text{L}} + \Lambda_1 - \sqrt[3]{\frac{\Delta_1}{36}}\left[\left(9 + f_3^2 + f_4^2 + f_6^2 + f_7^2\right)\mathbf{1}\right. \tag{7.32}$$
$$\left. +2f_3f_4Z_0Z_1 - 2Z_0Z_2 + 4Z_3\right] + O\left(\Delta_1^{-2}\right).$$

### 7.5.2 Third Order

The target Hamiltonian terms appears at third order from processes that transition between degenerate ground states. However, there is also an additional, unwanted process which occurs at this order. This competing process involves one interaction with $H_{2\text{L}}$ and $\Lambda_1$ in the high-energy subspace,

$$\frac{1}{\Delta_1^2}V_{-+}V_+V_{+-}^{(1)} = \frac{\mu_1^2}{\Delta_1^2}[A_1\left(H_{2\text{L}} + \Lambda_1\right)A_1 + B_1\left(H_{2\text{L}} + \Lambda_1\right)B_1 + C_1\left(H_{2\text{L}} + \Lambda_1\right)C_1 \tag{7.33}$$
$$+ A_2\left(H_{2\text{L}} + \Lambda_1\right)A_2 + B_2\left(H_{2\text{L}} + \Lambda_1\right)B_2 + C_2\left(H_{2\text{L}} + \Lambda_1\right)C_2 + A_3\left(H_{2\text{L}} + \Lambda_1\right)A_3$$
$$+ B_3\left(H_{2\text{L}} + \Lambda_1\right)B_3 + C_3\left(H_{2\text{L}} + \Lambda_1\right)C_3].$$

These processes are illustrated diagrammatically in Figure 7.6.

The process we want occurs with the ancilla transition $|000\rangle \rightleftharpoons |111\rangle$ which flips all three bits (for each term separately since they have different ancillae). There are $3! = 6$

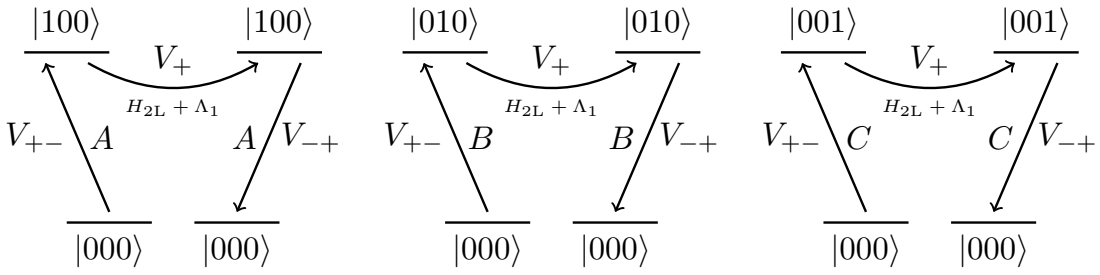



Figure 7.6: Diagrams for the competing process encountered at third order. Note that each of these diagrams can also occur backwards if the system starts in $|111\rangle$.

possible ways to flip the bits for each term, (these processes are illustrated in Figure 7.3),

$$\begin{aligned}\frac{1}{\Delta_1^2}V_{-+}V_{+}V_{+-}^{(2)} &= 6\frac{\mu_1^3}{\Delta_1^2}\left(A_1B_1C_1 + A_2B_2C_2 + A_3B_3C_3\right)\\ &= A_1B_1C_1 + A_2B_2C_2 + A_3B_3C_3. \end{aligned} \tag{7.34}$$

Because $H_{2L}$ has no $\Delta_1$ dependence and $\mu_1$ is order $O\left(\Delta_1^{2/3}\right)$, terms such as $\left(\mu_1^2/\Delta_1^2\right)A_1H_{2L}A_1$ will vanish in the limit of large $\Delta_1$; therefore, the third order effective Hamiltonian is,

$$\begin{aligned} H_{\text{eff}}^{(3)} &= H_{2L} + \Lambda_1 - \sqrt[3]{\frac{\Delta_1}{36}}\left[\left(9 + f_3^2 + f_4^2 + f_6^2 + f_7^2\right)\mathbf{1}\right. \\ &+ \left. 2f_3f_4Z_0Z_1 - 2Z_0Z_2 + 4Z_3\right] + \frac{\mu_1^2}{\Delta_1^2}\left(A_1\Lambda_1A_1 + B_1\Lambda_1B_1\right. \\ &+ C_1\Lambda_1C_1 + A_2\Lambda_1A_2 + B_2\Lambda_1B_2 + C_2\Lambda_1C_2 + A_3\Lambda_1A_3 \\ &+ \left. B_3\Lambda_1B_3 + C_3\Lambda_1C_3\right) + A_1B_1C_1 + A_2B_2C_2 + A_3B_3C_3 \end{aligned} \tag{7.35}$$

with error $O\left(\Delta_1^{-3}\right)$. We see that if $\Lambda_1 = \frac{1}{\Delta_1}V_{-+}V_{+-}$ then the unwanted contribution at third order will go to zero in the limit of large $\Delta_1$ and the second order term will cancel exactly with $\Lambda_1$. Thus,

$$H_{\text{eff}}^{(3)} \approx H_{2L} + A_1B_1C_1 + A_2B_2C_2 + A_3B_3C_3 \tag{7.36}$$

$$H_{H_2} \to H_1 + V_1 \tag{7.37}$$

where "$\rightarrow$" denotes an embedding. There are still 3-local terms remaining in $V_1$,

$$\begin{aligned} V_1 &= \mu_1 \left(f_4 Z_0 + f_3 Z_1\right) X_{a_1} + \mu_1 X_2 \left(X_{b_1} + X_{b_2}\right) \\ &+ \mu_1 Z_3 X_{c_1} + \mu_1 f_7 Z_0 X_{a_2} + \mu_1 Z_1 \left(Z_{c_2} + X_{c_3}\right) + \mu_1 X_{b_3} \\ &+ \underbrace{\mu_1 Z_1}_{A_4} \underbrace{Z_3}_{B_4} \underbrace{\left(X_{c_2} + X_{c_3}\right)}_{C_4} + \underbrace{\mu_1 f_6 X_0}_{A_5} \underbrace{X_2}_{B_5} \underbrace{X_{a_3}}_{C_5} + \underbrace{\left(-\mu_1\right) Z_0}_{A_6} \underbrace{Z_2}_{B_6} \underbrace{X_{b_3}}_{C_6} + H_{2\mathrm{L}} + \Lambda_1. \end{aligned} \tag{7.38}$$

With this notation we reorganize our Hamiltonian a final time, so that $H_{\mathrm{H}_2} \rightarrow H_{2\mathrm{L}} + H_{3\mathrm{L}}$,

$$H_{3\mathrm{L}} = A_4 B_4 C_4 + A_5 B_5 C_5 + A_6 B_6 C_6 \tag{7.39}$$

$$\begin{aligned} H_{2\mathrm{L}} &= \left(f_0 + \frac{9\Delta_1}{4}\right) \mathbf{1} + f_2 Z_1 + f_3 Z_2 + f_4 Z_0 Z_2 + f_5 Z_1 Z_3 + f_1 Z_0 \left(\mathbf{1} + Z_1\right) \\ &- \frac{\Delta_1}{4} \left(Z_{a_1} Z_{b_1} + Z_{a_1} Z_{c_1} + Z_{b_1} Z_{c_1} + Z_{a_2} Z_{b_2} + Z_{a_2} Z_{c_2} + Z_{b_2} Z_{c_2} + Z_{a_3} Z_{b_3} + Z_{a_3} Z_{c_3}\right. \\ &\left. + Z_{b_3} Z_{c_3}\right) + \sqrt[3]{\frac{\Delta_1^2}{6}} \left[\left(f_4 Z_0 + f_3 Z_1\right) X_{a_1} + Z_3 X_{c_1} + f_7 Z_0 X_{a_2} + X_2 \left(X_{b_1} + X_{b_2}\right) + X_{b_3}\right. \\ &\left. + Z_1 \left(X_{c_2} + X_{c_3}\right)\right] + \sqrt[3]{\frac{\Delta_1}{36}} \left[\left(9 + f_3^2 + f_4^2 + f_6^2 + f_7^2\right) \mathbf{1} + 2 f_3 f_4 Z_0 Z_1 - 2 Z_0 Z_2 + 4 Z_3\right]. \end{aligned} \tag{7.40}$$

The third order gadget we need to reduce $H_{3\mathrm{L}}$ takes exactly the same form as before except with the term labels $1, 2, 3$ exchanged for the term labels $4, 5, 6$. The components of the final gadget are

$$\begin{aligned} H_2 &= \frac{9\Delta_2}{4} \mathbf{1} - \frac{\Delta_2}{4} \left(Z_{a_4} Z_{b_4} + Z_{a_4} Z_{c_4} + Z_{b_4} Z_{c_4}\right. \\ &\left. + Z_{a_5} Z_{b_5} + Z_{a_5} Z_{c_5} + Z_{b_5} Z_{c_5} + Z_{a_6} Z_{b_6} + Z_{a_6} Z_{c_6} + Z_{b_6} Z_{c_6}\right) \end{aligned} \tag{7.41}$$

and

$$\begin{aligned} V_2 &= \mu_2 \left(A_4 X_{a_4} + B_4 X_{b_4} + C_4 X_{c_4} + A_5 X_{a_5}\right. \\ &\left. + B_5 X_{b_5} + C_5 X_{c_5} + A_6 X_{a_6} + B_6 X_{b_6} + C_6 X_{c_6}\right) + H_{2\mathrm{L}} + \Lambda_2 \end{aligned} \tag{7.42}$$

where $\mu_2 = \sqrt[3]{\frac{\Delta_2^2}{6}}$ and

$$\begin{aligned}\Lambda_2 &= \frac{\mu_2^2}{\Delta_2}\left(A_4^2 + B_4^2 + C_4^2 + A_5^2 + B_5^2 + C_5^2 + A_6^2 + B_6^2 + C_6^2\right) \\ &= \sqrt[3]{\frac{\Delta_2}{6}}\left[\frac{7}{\sqrt[3]{6}} + \Delta_1^{4/3}\left(\frac{1}{3} + \frac{f_6^2}{6}\right)\right]\mathbf{1} + \sqrt[3]{\frac{2\Delta_2}{9}}X_{c_2}X_{c_3}. \end{aligned} \tag{7.43}$$

This time the spectral gap and Hamiltonian have the subscript "2" to associate them with our second application of perturbation theory. We have thus shown the embedding $H_{\mathrm{H}_2} \to H_2 + V_2$. We present an interaction graph for the embedded Hamiltonian in Figure 7.7.

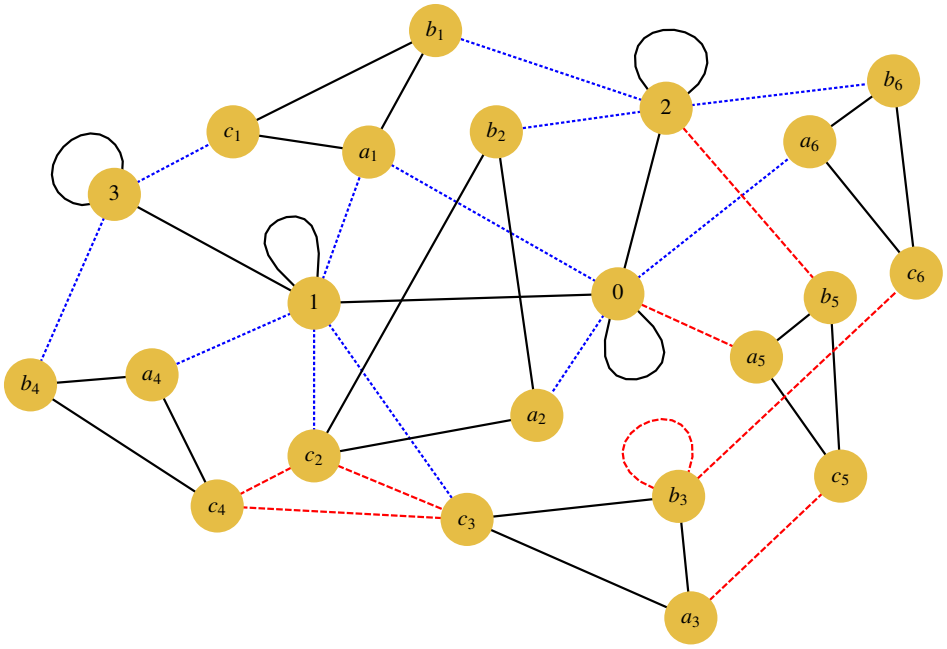



Figure 7.7: Interaction graph for embedded molecular hydrogen Hamiltonian. Each node represents a qubit. The solid, black edges represent $ZZ$ terms and the black loops represent local $Z$ terms. The dashed, red edges represent $XX$ terms and the red loops represent local $X$ terms. The dotted, blue edges represent $XZ$ terms. It is easy to see the unperturbed Hamiltonians corresponding to the six 3-operator terms (the black triangles).

## 7.6 Conclusion

We have presented a fully general method for mapping any molecular electronic structure instance to a 2-local Hamiltonian containing only $ZZ$, $XX$ and $XZ$ terms. Our method is scalable in the sense that all experimental resources (qubits, control precision, graph degree) scale polynomially in the number of orbitals. We used perturbative gadgets which embed the entire target Hamiltonian (as opposed to just the ground state), thus guaranteeing that the eigenvalue gap is conserved under our reduction. Furthermore, we showed that bit-flip gadgets can be applied to remove experimentally challenging $YY$ terms. The resulting Hamiltonian is suitable for implementation in superconducting systems, quantum dots and other systems of artificial spins with the correct engineered interactions.

Further reduction of the types of interactions present is possible, to either $ZZ$ and $XX$ terms or $ZZ$ and $XZ$ terms, using the techniques of [38]. This makes the required interactions for simulating electronic structure Hamiltonians equivalent to the requirements of universal adiabatic quantum computation [38]. However, repeated reduction of the Hamiltonian results in more stringent precision requirements. The chosen target set of interactions strikes a balance between control precision and a reasonable set of distinct types of controllable interaction. The techniques developed here could also be applied to interacting fermion problems on the lattice. However, in that case it is possible to improve beyond the Bravyi-Kitaev mapping and exploit the locality of the interactions to directly obtain Hamiltonians whose locality is independent of the number of orbitals [274].

We intend to follow-up this work with an analysis of hardware requirements for implementation on a system with superconducting qubits. A detailed scaling study of the exact resources needed for this algorithm as a function of molecular size is underway. We propose to read out energy eigenvalues using the tunneling spectroscopy of a probe qubit.

This technique has already been demonstrated experimentally with rf SQUID flux qubits in [29]. In this scheme, a probe qubit is coupled to a single qubit of the simulation. Tunneling transitions allow the probe qubit to flip when the energy bias of the probe is close to an eigenvalue of the original system. Hence detection of these transitions reveals the eigenspectrum of the original system. In this way, we would be able to directly measure the eigenspectra of the molecular systems embedded into the spin Hamiltonian using the techniques developed in the present paper.

There has been rapid recent progress in new classical algorithms, such as DMRG (density matrix renormalization group) and related tensor network methods, and proving complexity and approximability results pertaining to minimal resource model Hamiltonians. By using and understanding the techniques we have introduced in this paper, problems in chemistry can be reduced to such models and these discoveries can be leveraged to make advances in electronic structure theory. However, we note that the spin Hamiltonians that result from the mapping developed here will be *non-stoquastic*, and classical simulation techniques will therefore suffer from the fermionic sign problem [55]. This further motivates the construction of quantum hardware to address the electronic structure problem by quantum simulation of these spin Hamiltonians.

## Acknowledgements

The authors thank Sarah Mostame for helpful comments during revision. R.B. thanks Yudong Cao and Jacob Biamonte for many helpful conversations about perturbative gadgets. A. A.-G. acknowledges funding from National Science Foundation grant CHE-1152291 and the Air Force Office of Scientific Research under contract FA9550-12-1-0046. P.J.L. acknowledges National Science Foundation grant PHY-0955518. A.A.-G. and R. B. thank the

Corning Foundation for their generous support. Research sponsored by the United States Department of Defense. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressly or implied, of the United States Government.

# Part III

# Digital Quantum Computation

## Chapter 8

# Scaling of Trotter-Suzuki Errors in Quantum Chemistry Simulation

Apart from minor modifications, this chapter originally appeared as [15]:

"Chemical Basis of Trotter-Suzuki Errors in Quantum Chemistry Simulation". Ryan Babbush, Jarrod McClean, Dave Wecker, Alán Aspuru-Guzik and Nathan Wiebe. *Physical Review A*. Volume 91, Number 2: 022311. 2015.

## Abstract

Although the simulation of quantum chemistry is one of the most anticipated applications of quantum computing, the scaling of known upper bounds on the complexity of these algorithms is daunting. Prior work has bounded errors due to Trotterization in terms of the norm of the error operator and analyzed scaling with respect to the number of spin orbitals. However, we find that these error bounds can be loose by up to sixteen orders of magnitude for some molecules. Furthermore, numerical results for small systems fail to reveal any clear correlation between ground state error and number of spin orbitals. We instead argue that chemical properties, such as the maximum nuclear charge in a molecule and the filling fraction of orbitals, can be decisive for determining the cost of a quantum simulation. Our analysis motivates several strategies to use classical processing to further reduce the required

Trotter step size and to estimate the necessary number of steps, without requiring additional quantum resources. Finally, we demonstrate improved methods for state preparation techniques which are asymptotically superior to proposals in the simulation literature.

## 8.1 Introduction

The idea that the simulation of quantum systems would be efficient on a quantum computer dates back to Feynman's original work on quantum mechanical computers [107]. Almost a decade after Abrams and Lloyd [1] demonstrated a scalable scheme for the quantum simulation of fermions, Aspuru-Guzik *et al.* [11] proposed that these techniques could be used to efficiently determine the ground state energy of molecular Hamiltonians, solving what chemists refer to as the electronic structure problem. Since then, a great deal of work has focused on specific strategies for the quantum simulation of quantum chemistry. While most of these approaches are based on a second quantized representation of the problem making use of both phase estimation and Trotterization [217, 11, 277, 276, 285, 77, 243, 281, 140, 224, 195], recently some have proposed alternative schemes such as the quantum variational eigensolver [223], an adiabiatic algorithm [14] and an oracular approach based on a 1-sparse decomposition of the configuration interaction Hamiltonian [264]. In fact, quantum chemistry is such a popular application that toy problems in chemistry have been solved on a variety of experimental quantum information processors which include quantum optical systems [175, 223], nuclear magnetic resonance [97] and solid-state Nitrogen-vacancy center systems [279].

Recently, a series of papers [281, 140, 224, 195] has provided improved analytical and empirical bounds on the resources required to simulate classically intractable benchmarks using a quantum computer. While the initial findings in [281] were pessimistic, improve-

ments in both bounds and algorithms introduced in [140] and [224] have reduced these estimates by more than thirteen orders of magnitude for simulations of Ferredoxin. The primary contribution of [195] was to point out that in the limit of large molecules, the use of a local basis can substantially reduce asymptotic complexity of these algorithms. In this paper we build on the findings of [281, 140, 224, 195] to offer new perspectives regarding the scaling of the second quantized, Trotterized, phase estimation algorithm for quantum chemistry. In particular, we question a basic assumption implicit in all of these works: that the Trotter error explicitly depends on the number of spin orbitals being simulated.

Instead, we argue that chemical properties such as the filling fraction of electrons in a given basis, the particular choice of orbital basis and the nuclear potential play a more significant role in determining the Trotter error than does the number of spin orbitals for small molecules. We support these arguments with numerical analysis based on the explicit computation of the Trotter error operator derived in [224]. Additionally, we show that classically tractable approximations to the ground state wavefunction can be used to efficiently estimate the Trotter error expected in a particular ground state simulation. This result is of significant practical importance because without a procedure for estimating the Trotter error, one must rely on analytical error bounds which (as we show) tend to overestimate the ground state error by many orders of magnitude. Finally, we show asymptotically improved circuits for state preparation based on these classical ansatz states.

### 8.1.1 The electronic structure problem

The electronic structure problem is to estimate the energy of electrons interacting in a fixed nuclear potential to within an additive error of $\epsilon$. This Hamiltonian may be written

as,

$$H = -\sum_i \frac{\nabla_{r_i}^2}{2} - \sum_{i,j} \frac{Z_i}{|R_i - r_j|} + \sum_{i,j>i} \frac{1}{|r_i - r_j|} \tag{8.1}$$

where we have used atomic units, $\{R_i\}$ denotes nuclear coordinates, $\{r_i\}$ electronic coordinates, and $\{Z_i\}$ nuclear charge. Often times, the utility of these energies is to provide Born-Oppenheimer surfaces for molecular modeling at finite temperatures. Usually, chemists are interested in obtaining free energy landscapes which provide mechanistic insight into chemical events of significant practical importance such as drug binding, catalysis and material properties. These free energy landscapes must be extremely accurate as chemical rates are exponentially sensitive to changes in free energy. Under typical laboratory conditions of room temperature and atmospheric pressure, "chemical accuracy" is required which sets $\epsilon$ to the order of $10^{-3}$ hartree [145] where 1 hartree is $\frac{\hbar^2}{m_e e^2 a_0^2}$ and $m_e$, $e$ and $a_0$ denote the mass of an electron, charge of an electron and Bohr radius, respectively.

We represent the electronic structure Hamiltonian in second quantization [145] as this requires significantly fewer qubits than approaches using the first quantized Hamiltonian [298, 162],

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s \tag{8.2}$$

in which creation and annihilation operators act on a basis of orthogonal spin orbitals, $\{\varphi_i\}$ and the one-electron and two-electron integrals are

$$h_{pq} = \int d\sigma \ \varphi_p^*(\sigma) \left( -\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \varphi_q(\sigma) \tag{8.3}$$

$$h_{pqrs} = \int d\sigma_1 \ d\sigma_2 \ \frac{\varphi_p^*(\sigma_1)\varphi_q^*(\sigma_2)\varphi_s(\sigma_1)\varphi_r(\sigma_2)}{|r_1 - r_2|} \tag{8.4}$$

where $\sigma_i$ contains spatial and spin degrees of freedom for the electrons. The operators $a_p^\dagger$

and $a_r$ obey the fermionic anti-commutation relations

$$\{a_p^\dagger, a_r\} = \delta_{p,r}, \qquad \{a_p^\dagger, a_r^\dagger\} = \{a_p, a_r\} = 0. \tag{8.5}$$

In principle, the number of spin orbitals used to represent a molecule is not a property of the molecule. However, the quantum chemistry community has certain conventions (based on periodic trends) for the number of spin orbitals that should be used for each atom in the period table, depending on the desired level accuracy in the calculation. In a minimal basis, first period atoms receive two spin orbitals, second period atoms receive 10 spin orbitals and third period atoms receive 18 spin orbitals. The reasoning behind this scheme is that the most important orbitals are those which have a principal quantum number less than or equal to that of the highest occupied orbital according to Hund's rules.

In addition to choosing a spatial basis, one must choose an orbital basis that associates orthogonal spatial functions constructed from the spatial basis, with the second quantized sites. Throughout this paper we investigate three such orbital basis sets: the "local basis" is the set of orthogonal atomic orbitals discussed in [195], the "canonical basis" is the Hartree-Fock molecular orbitals, and the "natural basis" is that which diagonalizes the one-electron density matrices associated with the exact ground state[1]. It is worth pointing out that the canonical orbitals are the natural orbitals of a Hartree-Fock calculation using a single determinant.

From Eq. (8.2), we see that the number of terms in the Hamiltonian scales as $\Theta\left(N^4\right)^2$. However, McClean *et al.* [195] recently pointed out that the basis functions decay super-

[1]The natural basis can be well approximated without performing an exact calculation by repeating truncated configuration interaction calculations from reference states defined using the natural orbitals associated with a previous solution.

[2]We use the typical computer science convention that $f \in \Theta(g)$, for any functions $f$ and $g$, if $f$ is asymptotically upper and lower bounded by a multiple of $g$, $\mathcal{O}$ indicates an asymptotic upper bound, $\tilde{\mathcal{O}}$ indicates an asymptotic upper bound up to polylogarithmic factors, $\Omega$ indicates the asymptotic lower bound and $f \in o(g)$ implies $f/g \to 0$ in the asymptotic limit.

exponentially with distance in a local basis. This means that the integrals in Eq. (8.3) and Eq. (8.4) will be negligibly small for many of the orbitals which in turn allows the number of terms in the Hamiltonian to be truncated to $\tilde{\mathcal{O}}\left(N^2\right)$ or $\tilde{\mathcal{O}}\left(N\right)$ depending on the size and geometry of the molecule. All of the particular benchmarks studied in this paper involve less than four atoms and so we consider the number of non-negligible terms in the Hamiltonian to scale as $\Theta(N^4)$, even in a local basis.

### 8.1.2 Quantum simulation of quantum chemistry

The electronic structure problem is classically intractable to current methods even after discretizing the Hilbert space. This intractability can be understood as a consequence of the exponential size of the Hilbert space for the second quantized Hamiltonian. Similarly, existing methods such as configuration interaction, require consideration of a number of electronic configuration states that increases exponentially as the approximation becomes more exact. Quantum simulation offers a way to circumvent these challenges by directly mapping the chemical system onto a set of qubits that can be manipulated using a quantum computer. The particular problem that we focus on is the problem of computing the ground state energy of the system. Other important physical quantities such as dipole moments can be found by evaluating their expectation value with respect to the prepared state. The simulation problem that we consider is as follows.

**Problem:** *Assume that the user is provided with a classical database containing $h_{pq}$ and $h_{pqrs}$ for a molecule with $N$ spin orbitals and a blackbox state preparation algorithm that prepares an approximation $|\tilde{0}\rangle$ to the ground state $|0\rangle$ such that $|\langle\tilde{0}|0\rangle|^2 \in \Omega\left(poly\left(N^{-1}\right)\right)$. Design a quantum circuit that uses these elements to estimate the ground state energy of Eq. (8.2) within additive error $\epsilon$ using a minimal expected number of gates and qubits.*

Most proposals for quantum computer simulation of chemical systems use similar strategies to solve this problem. The first step involves translating the basis of the second quantized Hamiltonian to that of the quantum computer. The standard way to do this is to use the occupation number basis in which individual qubits encode the occupation of a spin orbital. For example, the state $|00011\rangle$ would refer to an electronic state where the first two spin orbitals are occupied and the remaining three spin orbitals are unoccupied.

Although representing states is trivial, representing the Hamiltonian is not. The reason is that, although it may seem that the creation and annihilation operators $a_i^\dagger$ and $a_i$ are translated to $(X_i - iY_i)/2$ and $(X_i + iY_i)/2$ respectively, the resulting operators do not obey the anti-commutation relations in Eq. (8.5). This problem is addressed by using either the Jordan-Wigner transformation [257, 11] or the Bravyi-Kitaev transformation [56, 243] to modify these operators to have the correct anti-commutation relations. Importantly, the operators that result from using either of these representations are tensor products of Pauli operations. While the number of such terms in the transformed Hamiltonian scales as $\mathcal{O}(N^4)$ using both approaches, the locality (i.e. many-body order) of these terms scales as $\mathcal{O}(N)$ under the Jordan-Wigner transformation and $\mathcal{O}(\log N)$ under the Bravyi-Kitaev transformation [243].

Since exponentials of a polynomial number of Pauli operators are known to be efficiently simulatable, $e^{-iHt}|\tilde{\psi}\rangle$ can be implemented using a polynomial number of gates using a quantum computer. There are many different approaches that can be used to achieve this and the majority of these rely on Trotter decompositions, which we will discuss in more detail later. However, each of these methods solves a dynamical simulation problem and does not directly solve the ground state energy estimation problem. The phase estimation algorithm provides the connection needed to relate the eigenvalue estimation problem to a

dynamical simulation problem.

The quantum phase estimation algorithm (PEA) uses a quantum computer to efficiently estimate energies from the phases $\{\theta_n(t)\}$ accumulated during time evolution under a propagator $U_H(t)$ associated with the Hamiltonian of interest $H$; i.e.

$$e^{iHt}|n\rangle = U_H(t)|\psi_n\rangle = e^{i\theta_n(t)}|\psi_n\rangle \tag{8.6}$$

$$\theta_n(t) = (E_n t) \mod 2\pi \tag{8.7}$$

where $\{|\psi_n\rangle\}$ and $\{E_n\}$ represent eigenstates and eigenvalues of $H$. If we initialize a quantum register in a state $|\tilde{\psi_0}\rangle$ then time evolution under a static Hamiltonian produces the superposition,

$$U_H(t)|\tilde{\psi_0}\rangle = \left(\sum_{n=0}^{2^N-1} e^{i\theta_n(t)}|\psi_n\rangle\langle\psi_n|\right)|\tilde{\psi_0}\rangle\,. \tag{8.8}$$

Measuring the phase of this superposition projects the system to state $|\psi_0\rangle$ with probability $|\langle\psi_0|\tilde{\psi_0}\rangle|^2$. Thus, under the assumptions of our problem, at most a polynomial number of repetitions of the phase estimation algorithm will be needed to find the ground state energy.

There are obviously two contributions to the cost of solving the electronic structure problem via quantum computing: ($a$) the overlap $|\langle\psi_0|\tilde{\psi_0}\rangle|^2$ and (b) the cost of simulating the dynamics of the system. Since the overlap is independent of the simulation method used (to second-order in perturbation theory) most work on the topic has focused on reducing the latter cost. We discuss both of these issues in the following.

Our main focus is on Trotter-Suzuki based methods, which involve a discretization of the time evolution known as Trotterization. Trotterization approximates $U_H(t)$ as a series of time steps known as "Trotter slices" during which only one of the Hamiltonian terms is actually active. A Trotter series containing $\mu$ Trotter slices is said to have a "Trotter number" of $\mu$ and the error in this approximation, which arises from non-commutativity of

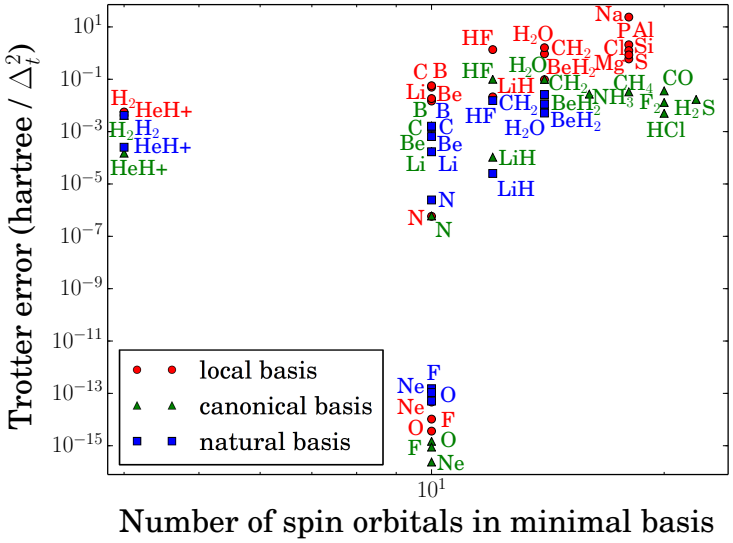



Figure 8.1: Spin-orbitals versus ground state Trotter error for various molecular benchmarks in three different basis sets. Despite analytical predictions to the contrary (in prior works), it would appear that no clear relation holds between the Trotter error induced on the ground state and the number of spin orbitals for these benchmarks.

the Hamiltonian terms, vanishes as $\mu \to \infty$. For a fixed order Trotter-Suzuki formula, each Trotter slice contains a number of gates that is proportional to the number of terms in the Hamiltonian, $m$. The value of $m$ depends on basis and molecular size and its scaling with $N$ ranges from $\tilde{\mathcal{O}}(N) - \tilde{\mathcal{O}}(N^4)$. Since the the total complexity of the quantum simulation circuit for chemistry is $\tilde{\mathcal{O}}(m\mu)$, understanding how both of these terms scales is vital for determining whether quantum chemistry will be viable on small scale quantum computers.

The big question that several recent papers have attempted to address is: "how does $\mu$ scale with $N$?" Indeed, this issue is central to the optimizations introduced in many of these simulation methods. Given the importance of this issue in the literature, the data in Figure 8.1 may come as a complete surprise. We see there that for modestly small molecules, the error in the second-order Trotter-Suzuki formula does not have a clear functional dependence on $N$. This is especially surprising for cases of canonical and natural

orbitals where there is little evidence of even an increasing trend in the error as a function of $N$. This lack of monotonicity is particularly striking for the atoms N, O, F, Ne which show negligibly small Trotter errors. In fact, for these molecules (along with others such as Helium Hydride and Lithium Hydride) $\mu = 1$ or $\mu = 2$ is sufficient to achieve chemical accuracy despite the fact that their Hamiltonians contain hundreds of non-commuting terms.

In order to understand why the Trotter error deviates so strongly from prior expectations, we analyze a leading order perturbative expression for the error in the second-order Trotter formula. The insights gained from this analysis raise an interesting point: although there is not a strong correlation between $N$ and the Trotter error, other chemical properties play a decisive role in the Trotter error. This forces us to reconsider how we conceptualize the scaling of quantum chemistry simulation relative to prior results in quantum simulation, e.g. [1, 11, 162, 4, 71, 34, 35, 140, 224, 195].

## 8.2 Analysis of Trotter error operator

The second-order Trotter-Suzuki decomposition allows us to approximate the propagator as a series of unitaries corresponding to the individual Hamiltonian terms. In particular, the second-order[3] Trotter formula gives us,

$$U_H^{\mathrm{TS}}(\Delta_t) \equiv \prod_{\alpha=0}^{m-1} U_{m-\alpha}\left(\frac{\Delta_t}{2}\right) \prod_{\alpha=1}^{m} U_\alpha\left(\frac{\Delta_t}{2}\right) \tag{8.9}$$

where,

$$U_\alpha\left(\frac{\Delta_t}{2}\right) = e^{-iH_\alpha \Delta_t/2}. \tag{8.10}$$

[3]Note that in work that focuses on high-order Trotter-Suzuki formulas Eq. (8.9) is often called the first-order Trotter Suzuki formula because it is the lowest iteration order in Suzuki's iterative construction of high-order splitting formulas.

The second-order formula applies each unitary twice with the second half of the Trotter series in reverse order of the first half to cancel out error terms in the ground state energy that would arise at first-order in $\Delta_t$. We use this to make the approximation, valid for sufficiently small values of $\Delta_t$, that

$$U = e^{iHt} \approx \left(U^{\mathrm{TS}}\left(\Delta_t\right)\right)^{\mu}, \quad \Delta_t = t/\mu. \tag{8.11}$$

Poulin *et al.* [224] focus on bounding the error in this approximation with the Baker-Campbell-Hausdorff (BCH) formula,

$$\log\left(e^X e^Y\right) = X + Y + \frac{1}{2}\left[X, Y\right] + \frac{1}{12}\left[X, \left[X, Y\right]\right] - \frac{1}{12}\left[Y, \left[X, Y\right]\right] + ... \tag{8.12}$$

By recursively applying Eq. (8.12) to Eq. (8.9), the error operator may be written as $V = \sum_{j=1}^{\infty} V^{(j)}$. The leading order term in this expansion is,

$$V^{(1)} = -\frac{\Delta_t^2}{12} \sum_{\alpha \leq \beta} \sum_{\beta} \sum_{\gamma < \beta} \left[H_\alpha \left(1 - \frac{\delta_{\alpha,\beta}}{2}\right), \left[H_\beta, H_\gamma\right]\right] \tag{8.13}$$

with errors on the order of $\mathcal{O}\left(\Delta_t^4\right)$.

The leading order shift in the energy of the $i^{\mathrm{th}}$ eigenstate is given by non-degenerate perturbation theory as,

$$\Delta E_i = \left\langle \psi_i \right| V^{(1)} \left| \psi_i \right\rangle + \mathcal{O}\left(\Delta_t^4\right) \tag{8.14}$$

where $H\left|\psi_i\right\rangle = E_i \left|\psi_i\right\rangle$. Solving the electronic structure problem requires fixed precision in the energy, i.e. $\Delta E = \mathcal{O}\left(1\right)$. This suggests that we must shrink the time step for larger problem instances in order to offset any increase in Trotter error. In order to make the leading order shift in the energy eigenvalue at most $\delta$ it suffices to take

$$\mu = \mathcal{O}\left(t \sqrt{\frac{1}{\delta}\left\langle \sum_{\alpha \leq \beta} \sum_{\beta} \sum_{\gamma < \beta} \left[H_\alpha \left(1 - \frac{\delta_{\alpha,\beta}}{2}\right), \left[H_\beta, H_\gamma\right]\right]\right\rangle}\right). \tag{8.15}$$

Higher-order Trotter-Suzuki algorithms can be used to reduce the scaling of $\mu$; however they require a number of gates that scales exponentially with the order of the Trotter formula. This means that for many problems with modest error tolerances, the second-order Trotter formula Eq. (8.9) yields the most efficient results. Although a similar expression based on degenerate perturbation theory must be used for molecules near disassociation, in most practical cases Eq. (8.15) will accurately predict the required Trotter number in the limit of small $\delta$.

In practice, it is difficult to determine precisely how this error scales with problem size for real molecules. By inspection of Eq. (8.13), a loose bound of $\mu = \mathcal{O}\left(N^5\right)$ is obtained [224]. This bound is obtained by recognizing that the double commutator sum in Eq. (8.13) contains $\mathcal{O}(N^{12})$ terms but only $\mathcal{O}(N^{10})$ such terms are non–zero. In some cases, such as large molecules represented in a local orbital basis, many of these interactions can be neglected and the actual scaling of $\mu$ needed to achieve chemical accuracy may be closer to $\mu = \tilde{\mathcal{O}}\left(N^3\right)$ or $\mu = \tilde{\mathcal{O}}\left(N^{3/2}\right)$.

All of these scalings follow from worst case assumptions about the error and liberal application of the triangle inequality. Such arguments are not sufficient to explain the data in Figure 8.1 which does not show a clear dependence of $\mu$ on $N$. We therefore focus in the remainder on two quantities: $(a)$ the error in the ground state energy and $(b)$ the operator norm of the Trotter error operator. While $(a)$ is the best measure of the error in quantum chemistry simulation, we also focus on $(b)$ because it upper bounds $(a)$ and because it can be well approximated without diagonalizing the Hamiltonian.

In the numerics that follow we construct error operators by explicitly computing all $\mathcal{O}(N^{10})$ nonzero terms in Eq. (8.13). Once all the terms in the error operator are constructed, we simplify the resulting expression by normal-ordering the result. Here normal-

ordering refers to a sorting process where any chain of creation and annihilation operators that result from Eq. (8.13) are reordered such that creation operators always occur at the left-most part of the chain. This reordering is done by using the anti-commutation relations in Eq. (8.5). For example, $a_2 a_1 a_1^\dagger a_3^\dagger = a_1^\dagger a_3^\dagger a_1 a_2 - a_3^\dagger a_2$. These normal-ordered terms are then grouped, allowing their actions on computational basis states to be efficiently computed.

The Trotter scheme we investigate does not use the coalescing strategies introduced in [224], which would surely lead to even more error cancellation. We use a minimal spatial basis (STO-6G). The Trotter series is ordered in the "interleaving" scheme introduced in [140] and PQRS terms are ordered lexicographically. All molecular integrals in this work were calculated at equilibrium configurations using the GAMESS electronic structure package [240, 122]. While computing the error operator is efficient, evaluating the error operator on an eigenstate of the Hamiltonian cannot be performed in polynomial time on a classical computer. Due to the expensive nature of these calculations, we limit our investigation to benchmarks containing less than twenty spin orbitals. We study the scaling of the norm of the Trotter error operator as this quantity is the focus of analytical bounds introduced in [281] and [224]. Though the bounds in [224] are based on a upper bound for the operator norm of the error operator, here we use the exact value of $\|V^{(1)}\|$.

### 8.2.1 Comparison of norm of error operator and ground state error

An important question to ask is, "how does the error in the simulated ground state energy compare to that predicted by the norm of the error operator?" This is important for two reasons. The first reason is that there can be substantial cancellation in the sum implicit in Eq. (8.14). This effect is also discussed in [224]. The second reason is that

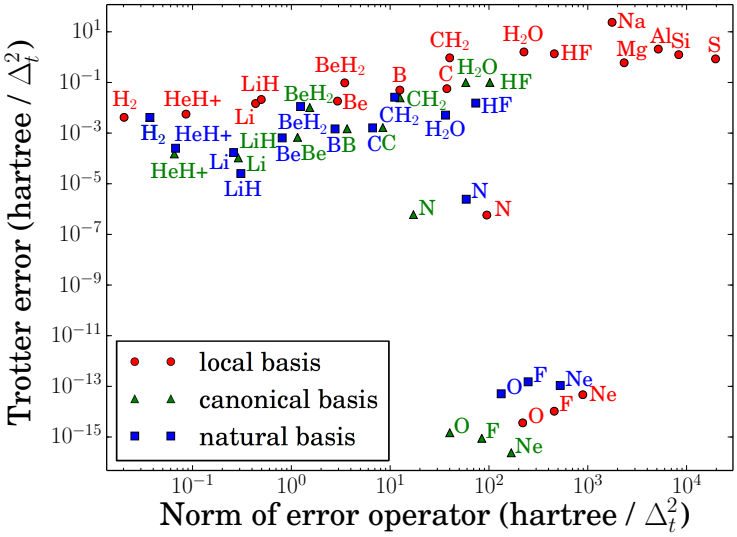



Figure 8.2: A comparison between the norm of the error operator and the error induced in the ground state. Notice that in many cases the basis of natural orbitals have the lowest Trotter error (especially for examples with large Trotter error).

the ground state may only have limited overlap with the eigenstates of the error operator that have large eigenvalues. We will discuss these two effects in detail later, but for now it suffices to ask how substantial the differences between the two measures are.

Figure 8.2 shows that substantial differences exist between the computed Trotter error and the norm of the error operator. In particular, for O, F and Ne these discrepancies can be as large as sixteen orders of magnitude. Other molecules, such as $H_2O$ and HF differ by only two orders of magnitude. This shows that existing estimates of the error can ludicrously overestimate the error in Trotter Suzuki formulas if the properties of the ground state are not also taken into account. Similar comparable results have also been observed for random many-body Hamiltonians [226].

To see this, let us consider Ne. By the convention for second-period atoms, Ne is given 10 spin orbitals in a minimal basis but it also has 10 electrons. This means that all of its spin orbitals will be occupied, i.e. $|\psi_0\rangle = |1\rangle^{\otimes 10}$. If we consider the action

of a single normal-ordered term from Eq. (8.14), $\alpha\, a_{p_1}^\dagger \cdots a_{p_5}^\dagger a_{q_1} \cdots a_{q_5}$, then we see that $\langle\psi_0|\, \alpha\, a_{p_1}^\dagger \cdots a_{p_5}^\dagger a_{q_1} \cdots a_{q_5}\, |\psi_0\rangle = 0$ unless $\{p_1, \ldots, p_5\} = \{q_1, \ldots, q_5\}$ up to permutations. Thus, the vast majority of the terms present in the error operator will evaluate to zero, irrespective of the magnitude of their coefficients. A similar argument can be made for F and O except that the ground state will no longer precisely be the Hartree-Fock state and instead will be a linear combination of computational basis states. Nonetheless, it is easy to see that the vast majority of these expectation values will be zero for these highly constrained systems. We therefore expect from this argument that molecules that have spin orbitals that are nearly fully occupied will have abnormally low error compared to molecules that are half filled where the dimension of the space is maximal for a given number of basis functions. This not only justifies the shockingly small error in N, O, F, and Ne but also explains why only considering the norm of the error operator obscures this trend.

For most benchmarks there is still evidence of correlation between the norm of the error operator and the Trotter error. This means that trends in the norm of the error operator are often reflected in the simulation error. As we have seen, the properties of the molecules in question can change the nature of this relationship.

### 8.2.2 Dependence on basis

In addition to showing that Trotter error in the ground state is usually substantially less than the error operator norm, Figure 8.2 suggests that the error is also basis dependent. While previous works have focused on the local and canonical basis sets, this figure suggests that using natural orbitals can often lower Trotter error by several orders of magnitude relative to a local orbital basis.

Furthermore, we argue that the discrepancy between error norm and ground state

Table 8.1: Ratio of ground state error to error operator norm for molecular hydrogen in various basis sets.

| Basis | Type | Orbitals | error / norm |
|---|---|---|---|
| STO-6G | local | 4 | 0.2063 |
| 3-21G | local | 8 | 0.0568 |
| 6-31G | local | 8 | 0.0592 |
| 6-31++G | local | 12 | 0.0328 |
| STO-6G | canonical | 4 | 0.1131 |
| 3-21G | canonical | 8 | 0.0231 |
| 6-31G | canonical | 8 | 0.0242 |
| 6-31++G | canonical | 12 | 0.0108 |
| STO-6G | natural | 4 | 0.1131 |
| 3-21G | natural | 8 | 0.0472 |
| 6-31G | natural | 8 | 0.0547 |
| 6-31++G | natural | 12 | 0.0194 |

error increases with the number of spin orbitals to such an extent that the former should not be used to make arguments about the asymptotic scaling of the latter. One can always add more spin orbitals to a molecular Hamiltonian but given a reasonable orbital basis, the ground state and physically meaningful excited states will have increasingly limited occupancy in high energy orbitals. In this context, the energy of an orbital is understood to mean the energy of a single electron occupying that orbital in the absence of other electrons (appropriate for atomic orbitals) or in the presence of the average density of all other electrons (appropriate for the canonical orbitals). Additionally, the natural orbital basis is known to have the property that states with an odd number of excitations from ground state reference often have negligible overlap with the exact ground state [26].

While the error operator will inevitably contain many terms involving excitations to and from these high energy spin orbitals, eigenstates of physical interest (e.g. the ground state) are superpositions of configurations which have a limited number of excitations. Accord-

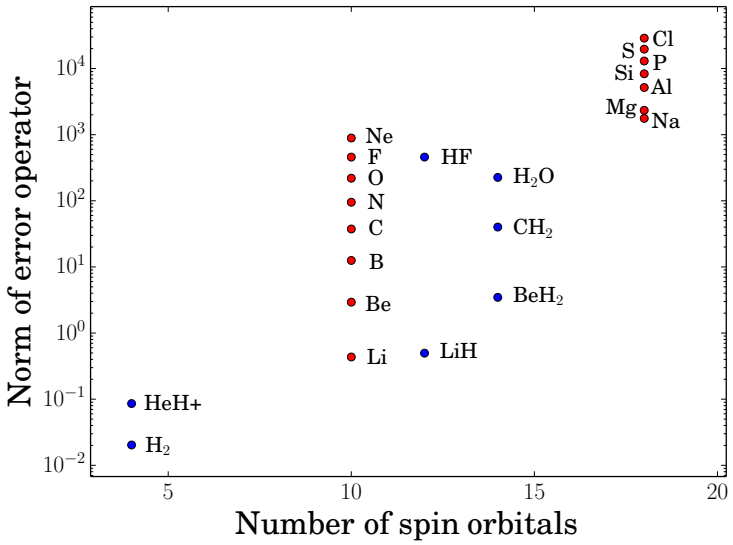



Figure 8.3: We correlate the number of spin orbitals with the norm of the error operator in the local basis. The semblance of a positive slope appears to be a symptom of increasing nuclear charge as the number of spin orbitals increase. Red dots are atoms and blue dots are molecules.

ingly, terms involving combinations of high energy orbitals are not expected to significantly contribute to the error induced in relevant eigenstates despite increasing the norm of the error operator. This principle is demonstrated in Table 8.1 which shows the ratio between ground state error and error norm for molecular hydrogen in various basis sets.

### 8.2.3 Dependence on nuclear charge

Figure 8.3 indicates that Trotter error norm correlates especially well with the maximum nuclear charge, as further demonstrated in Figure 8.4. The local basis is formed from the set of orthogonal atomic orbitals which are obtained for molecules using Löwdin symmetric orthogonalization on the original non-orthogonal local Gaussian orbitals [195]. These Gaussian basis functions are constructed as approximations to eigenfunctions of Hydrogen-like systems, with some fitting adjustments. As such, we can determine the scaling behavior

by considering the eigenfunctions of Hydrogen-like systems which are simple enough to permit analytical determination of how each term in the Hamiltonian will scale with nuclear charge. We begin by writing the eigenfunctions of a single electron in the potential of a point charge $Z$ in a convenient way,

$$\psi_{n\ell m}(\rho,\theta,\phi) = \sqrt{\left(\frac{2Z}{n}\right)^3 \frac{(n-\ell-1)!}{2n\,(n+\ell)!}} e^{-\frac{\rho}{n}} \left(\frac{2\rho}{n}\right)^{\ell} L_{n-\ell-1}^{2\ell+1}\left(\frac{2\rho}{n}\right) Y_{\ell}^{m}(\theta,\phi) \tag{8.16}$$

where $\rho = rZ$, $L_{n-\ell-1}^{2\ell+1}\left(\frac{2\rho}{n}\right)$ is a generalized Laguerre polynomial of degree $n-\ell-1$, and $Y_{\ell}^{m}(\theta,\phi)$ is a spherical harmonic of degree $\ell$ and order $m$. With the convention,

$$\varphi_p(\sigma_i) = \psi_p(\rho_i,\theta_i,\phi_i)\,\chi(s_i) \propto Z^{3/2} \tag{8.17}$$

$$d\sigma_i = \frac{\rho_i^2\, d\rho_i}{Z^3} \sin(\theta_i)\; d\theta_i\, d\phi_i\, ds_i \propto Z^{-3} \tag{8.18}$$

$$\nabla^2 = Z^2\left(\frac{\partial^2}{\partial\rho^2} + \frac{2}{\rho}\frac{\partial}{\partial\rho}\right) + \frac{Z^2}{\rho^2\sin^2(\theta)}\frac{\partial^2}{\partial\phi^2} + \frac{Z^2}{\rho^2\sin^2(\theta)}\frac{\partial^2}{\partial\phi^2} \propto Z^2 \tag{8.19}$$

where $\chi(s_i)$ is a spin assignment and $\sigma$ represents all degrees of freedom for an electron, we rewrite Eq. (8.3) and Eq. (8.4) in terms of $\rho$, assuming a single nuclei,

$$h_{pq} = \int d\sigma\; \varphi_p^*(\sigma)\left(-\frac{\nabla^2}{2} - \frac{Z^2}{\rho}\right)\varphi_q(\sigma) \tag{8.20}$$

$$h_{pqrs} = \int d\sigma_1\, d\sigma_2 \frac{\varphi_p(\sigma_1)\,\varphi_q(\sigma_2)\,\varphi_s(\sigma_1)\,\varphi_r(\sigma_2)}{|\rho_1-\rho_2|/Z}. \tag{8.21}$$

For both integrals, factors of $Z$ from the differential volume elements $d\sigma$ cancel with factors of $Z$ from the spin orbitals $\varphi$ and we find that,

$$|h_{pq}| = \Theta\left(Z^2\right) \tag{8.22}$$

$$|h_{pqrs}| = \Theta(Z). \tag{8.23}$$

Thus, it is clear that we can upper bound the scaling of individual Hamiltonian terms with nuclear charge as $\mathcal{O}\left(Z_{\max}^2\right)$. While this result is rigorous only when the orbital basis is the basis of true atomic orbitals, we expect qualitatively similar behavior in other bases.

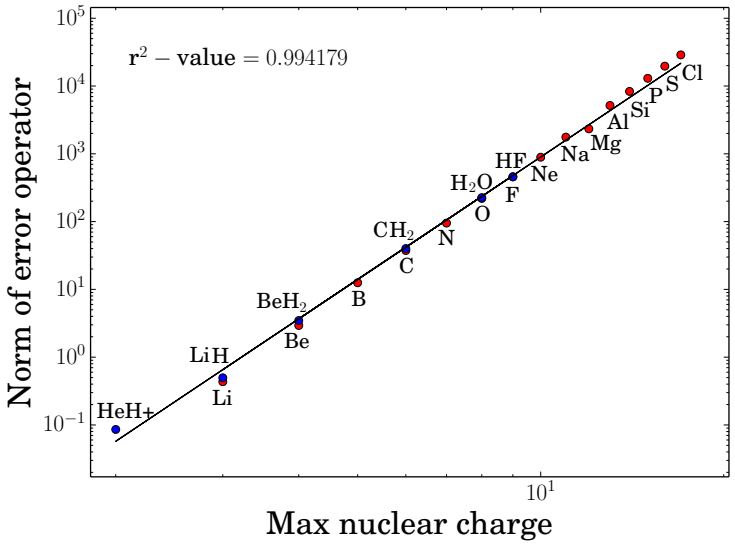



Figure 8.4: The norm of the error operator appears extremely well correlated with the maximum nuclear charge in a molecule when using a local basis of atomic orbitals. The black line is the line of best fit for a $Z_{\text{max}}^6$ scaling.

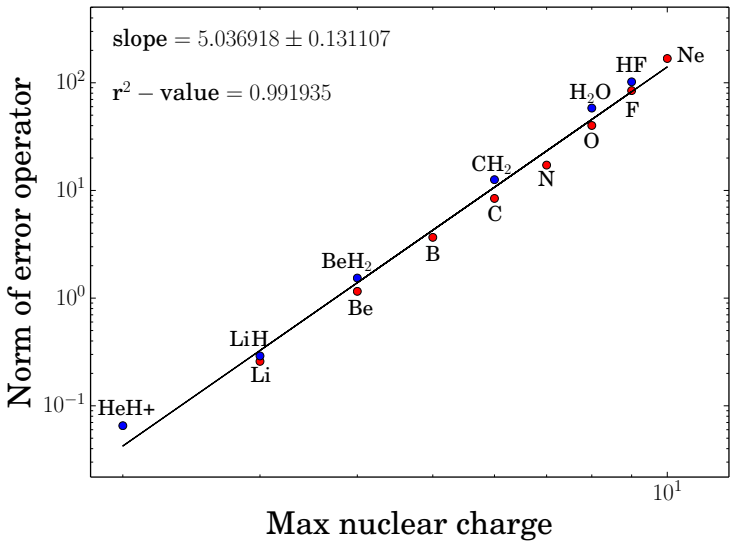



Figure 8.5: The norm of the error operator appears also well correlated with the maximum nuclear charge in a molecule when using the canonical basis of molecular orbitals. The black line is a least squares fit to the data which is roughly consistent with a $Z_{\text{max}}^5$ scaling.

Assuming the $h_{pq}$ terms dominates the error in the Trotter formula then Eq. (8.13) implies that the Trotter error should scale as $\mathcal{O}(Z_{\max}^6)$. This scaling is qualitatively consistent with the empirical scaling in Figure 8.4 which fits $Z_{\max}^6$ scaling to the norm of the error operator with an $r^2$-value of 0.994. Comparable results to this scaling have also been observed in diffusion Monte Carlo algorithms [65, 129].

These results imply that if an atomic basis is used then the error in the second-order Trotter-Suzuki formula scales at most as

$$\|V^{(1)}\| \in \mathcal{O}\left(N^4 Z_{\max}^6 + N^{10} Z_{\max}^3\right). \tag{8.24}$$

This result is a direct consequence of bounds on the Trotter-Suzuki error in [140] and the observation that double commutators of the one- and two-body terms produce at most $N^4$ and $N^{10}$ terms respectively. This implies that the computational complexity of performing the simulation on an arbitrary state, given fixed error tolerance of chemical accuracy, is $\mathcal{O}(N^4(N^2 Z_{\max}^3 + N^5 Z_{\max}^{3/2}))$. However, our numerical results are consistent with an $\mathcal{O}(N^4 Z_{\max}^3)$ which suggests that this scaling may be loose. It also important to note that the gate depth can be further reduced by using interleaving and nesting as per [140], which is significant when the algorithm is implemented on systems where quantum operations can be executed in parallel. It is also worth noting that the one-body terms dominate the two-body terms in every numerical example that we considered. Larger molecules with more $h_{pqrs}$ terms may lead to Trotter errors that scale as $\mathcal{O}(Z_{\max}^3)$ rather than $\mathcal{O}(Z_{\max}^6)$. More extensive numerical results may be needed to determine the conditions under which the two-body terms asymptotically dominate the one-body terms (if such conditions exist).

Figure 8.5 shows that these error estimates are pessimistic for the molecules considered when using the canonical basis. While the error norm is still strongly correlated to nuclear charge, unlike the scaling in the local basis, the fit to a $Z_{\max}^6$ scaling is less convincing.

Instead, the data empirically seems to follow a $Z_{\max}^5$ scaling. Intuitively, this is easy to envision because the molecular orbitals are inherently delocalized and thus it is natural to expect that the maximum nuclear charge should make less of an impact in this basis. We also see no evidence of explicit scaling with $N$ over this range in $Z_{\max}$. It is interesting to note that although the number of non-negligible integrals in a local orbital basis can be quadratically or quartically smaller than the size of an untruncated canonical molecular orbital basis, the scaling with $Z_{\max}$ seems to be better by a linear factor. This suggests interesting trade-offs between the two methods and hints that neither is intrinsically superior for quantum simulation.

### 8.2.4 Dependence on orbital structure

The terms that appear in the error operator include interactions between every orbital in the basis set. This begs the question of whether terms in the Hamiltonian that involve particular orbitals have larger contribution to the error. In order to assess this, we compute the error operator for a number of different molecules and normal-order the resultant operator. We then sum the magnitudes of every remaining term that either create or annihilate an electron in each of the orbitals. An example of this is provided in Figure 8.6, which shows the marginal coefficient magnitudes of all terms in the error operator (after normal-ordering) in terms of two spin orbitals they contain. Appendix 8.5.2 shows similar analysis for other molecules in other basis sets. As we can see, terms which involve the inner shell electrons dominate the norm of the error operator in the local basis.

We see from such figures that the inner orbitals, especially the single particle terms which are on the diagonal of the plot above, have a substantial impact on the Trotter error. This is not surprising as the inner atomic orbitals interact very strongly with nuclei so the

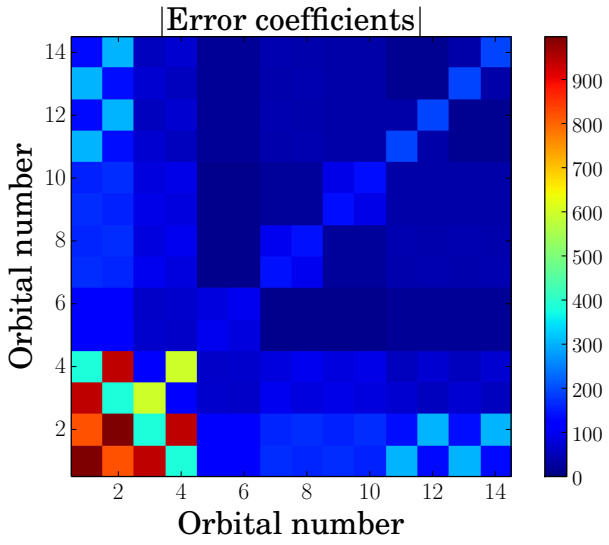



Figure 8.6: This plot shows the coefficients of normal-ordered terms in the error operator for water in a local basis as a function of the orbitals on which they act. The coefficients of the error terms are binned according to the orbitals involved in the term. This plot shows the marginal distribution of the magnitudes of those terms.

single particle integrals are likely to be much larger than the interaction integrals for these orbitals. Interestingly, although the valence shell electrons are often the most important for determining the chemical properties of a molecule, the inner orbitals are the ones that affect the error most significantly. This suggests that pseudo-potentials, which allow the core electrons to be treated as effectively "frozen", may provide a way to reduce the Trotter error in some circumstances. We leave this as an open question for future work.

### 8.2.5 Dependence on structure of eigenstates

Due to the substantial discrepancy between error induced on the ground state and operator norm, we might ask the following question: given the error operators for real molecules, what is the distribution of errors that would be induced on a random ensemble of vectors? This question is important as the answer will help us to identify the source

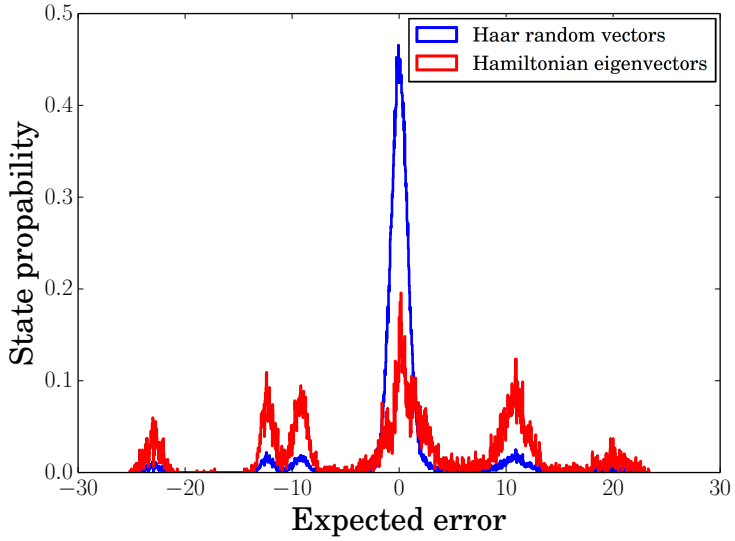



Figure 8.7: This plot shows the distribution of expectation values of the error operator for water in the local basis over its eigenstates and Haar random vectors. We see that the random vectors lead to substantially less error, on average, than do the Hamiltonian eigenstates. The Haar distribution of errors has a standard deviation of 4.82 while the Hamiltonian error distribution has standard deviation of 10.68.

of the observed error cancellation. We consider the ensemble of Haar random vectors which form a unitarily invariant ensemble of vectors with uniformly distributed complex elements. Unitary invariance ensures that the ensemble has uniform distribution in an arbitrary complete, orthonormal basis such as the eigenbasis of the error operator.

Denoting vectors from the random ensemble as $|v\rangle$ and eigenvectors of the error operator as $|k\rangle$ with eigenvalue $\lambda_k$, we are interested in analyzing properties of the following distribution of expected errors given by,

$$\Delta E\left(v\right) = \sum_k \lambda_k |\langle v\,|k\rangle|^2. \tag{8.25}$$

First, note that $\sum_k \lambda_k = 0$. This is because if $C = [A, B] = \sum_j \lambda_k \,|k\rangle\langle k|$ then

$$\sum_k \lambda_k = \mathrm{Tr(C)} = \mathrm{Tr(AB)} - \mathrm{Tr(BA)} = 0, \tag{8.26}$$

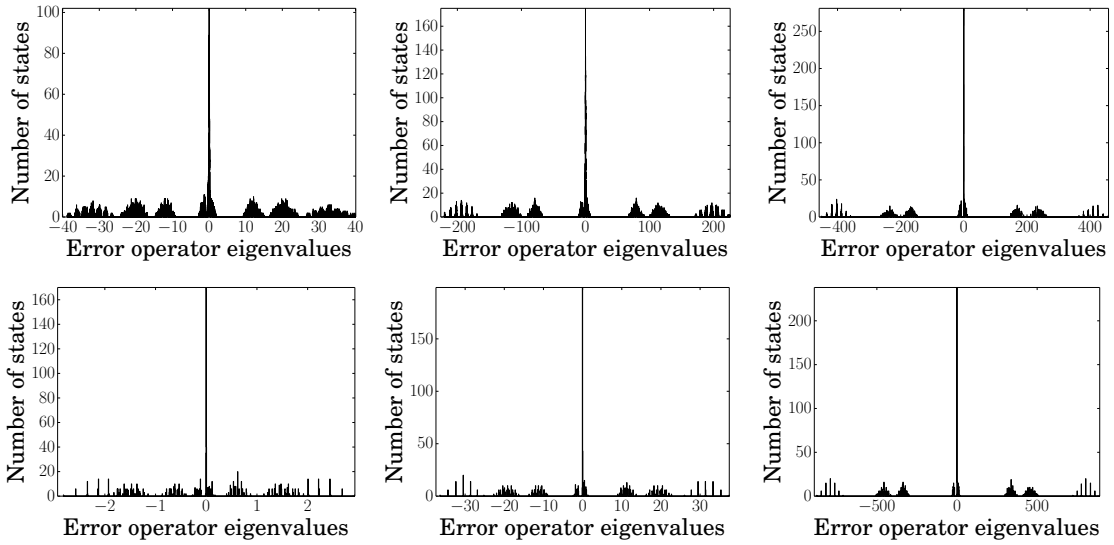



Figure 8.8: These are histograms of the eigenspecta of the error operators for various molecular and atomic benchmarks in the local basis. Proceeding clockwise from the top left, the molecules are water, Hydrogen Fluoride, Methylene, atomic Beryllium, atomic Carbon and atomic Neon. Error operators for all of our benchmarks have surprisingly similiar eigenspectra, regardless of the orbital basis. The source of this striking similarity and the reason for the particular structure is unknown.

from the cyclic property of the trace. Since $V^{(1)}$ is the sum of such operators, it follows that its trace is also zero. This implies that the Haar-expectation value of the error, over all possible random states, is

$$\mathbb{E}_H(\Delta E(v)) = \sum_k \lambda_k \mathbb{E}_H |\langle v\,|k\rangle|^2 = \frac{1}{2^N} \sum_k \lambda_k = 0. \tag{8.27}$$

This shows that there is no inherent bias that arises from Trotterization towards either overestimating or underestimating the true expectation value.

This result does not represent the *typical error* that we expect to see in a simulation. We also need to find the Haar variance of the expected error to estimate the typical variation of simulation errors about the mean. It is then easy to see that the Haar variance is

$$\mathbb{V}_H(\sum_k \lambda_k |\langle v|k\rangle|^2) = \sum_k \lambda_k^2 \mathbb{V}_H(|\langle v|k\rangle|^2). \tag{8.28}$$

In Appendix 8.5.1, we derive the Haar variance of the squared projection,

$$\mathbb{V}_H\left(|\langle v\,|k\rangle\,|^2\right) = \frac{2}{2^N\left(2^N+1\right)} - \frac{1}{2^{2N}}, \tag{8.29}$$

where $N$ is the number of spin orbitals. Combining Eq. (8.29) and Eq. (8.28) and using Chebyshev's inequality, we see that with high probability over $|v\rangle$

$$|\,\langle v|\,V^{(1)}\,|v\rangle\,| \in O\left(\frac{\sqrt{\sum_k \lambda_k^2}}{2^N}\right). \tag{8.30}$$

Eq. (8.30), surprisingly, shows that a concentration of measure argument causes the expectation of the Trotter error to be asymptotically zero if $(a)$ $|v\rangle$ is typical of a Haar random state, (b) $\sum_k \lambda_k^2 \in o(2^{2N})$ and $(c)$ $|v\rangle$ is chosen independently of the $|k\rangle$.

We do not expect a concentration of measure argument like this to hold for actual quantum simulations because it would imply that the Trotter errors *in eigenvalue estimation* shrink rapidly with system size for physically reasonable distributions of $\lambda_k$. Thus, it is natural to expect that one or both of assumptions $(a)$ and $(c)$ are not reasonable for eigenvalue estimation.

In Figure 8.7, we show the expected errors according to Eq. (8.25) over an ensemble of Haar random vectors as well as the expected errors over the eigenstates of the Hamiltonian for water. The results clearly show that the errors observed in this chemical example are much greater than we would expect from Haar random states. Furthermore, we see little evidence of concentration of measure of the errors about zero for the case where $|v\rangle$ is an eigenvector of $H$; whereas the Haar random $|v\rangle$ lead to results that are much more concentrated about zero error. This suggests that the discrepancies between the norm of the error operator and the ground state error cannot be explained by a simple randomization argument as the actual errors observed are much worse than would be otherwise expected.

Eq. (8.25) shows that the expected error is the convolution of the functions $\lambda_k$ and

$|\langle v|k\rangle|^2$. Thus, we expect the distribution of errors to resemble the underlying distribution of eigenvalues of $V^{(1)}$. This intuition can easily be seen by comparing Figure 8.7 to the eigenspectrum of the water error operator in Figure 8.8a. As expected, the distribution of errors for the random ensemble (Figure 8.7) resembles the error operator eigenspectrum (Figure 8.8a) with concentration about the mean (as anticipated by Eq. (8.30)). Also, it is interesting to note that the eigenspectra of the error operators for various molecules and atoms studied in this paper bear a remarkable degree of similarity and appear extremely structured as Figure 8.8 demonstrates. Additionally, every example has a sharp peak in its spectrum about zero error. This suggests that much of the rift between the norm of the error operator in Figure 8.2 may be due to the large number of eigenvectors with near-zero eigenvalue.

## 8.3 Improved simulation methods inspired by classical approaches

Given the large disparity between error operator norm and error induced on the exact ground state, any efficient method which allows one to approximate the error induced on the ground state (which implies an estimate for the number of Trotter steps needed) would be of critical importance for anyone wishing to actually run a quantum chemistry simulation on a quantum computer. A natural way to address this problem is to directly evaluate the error over a mesh in position and fit the data to a power law. This process can be made efficient using the SWAP test, as proposed by Wiebe et al [288]. A major drawback of this approach is that it requires roughly twice the qubits that the basic simulation used and also the variance in the estimate returned by the SWAP test can be prohibitively large. In this section, we propose an alternative method that estimates the error in the

ground state energy by evaluating the error operator on a classical ansatz for the ground state numerically. This method also allows the contribution to the error in the quantum simulation from the Trotter error to be subtracted off of the final estimate, improving the accuracy of the simulation without requiring additional quantum operations.

Perhaps the most well-known classical algorithm for solving the electronic structure problem is a mean-field approach known as the Hartree-Fock method [145]. In this scheme, single particle molecular orbitals are obtained using a self-consistent variational procedure in which each particle is made to interact with the average density of the other particles. The output of this calculation provides molecular orbitals which, together with a spin assignment, are used to approximate the $n$-particle wavefunction as an anti-symmetric product of the orbitals (known by chemists as a Slater determinant).

Unfortunately, the Hartree-Fock method is incapable of approximating dynamic electron correlation and is known to overestimate energies by an amount that is typically well above the threshold of chemical accuracy. To correct for this problem, one can expand the wavefunction in a basis of multiple Slater determinants and variationally solve for the coefficients which minimize the electronic energy. In general, there are $M = \binom{N}{n}$ valid configurations for $n$ electrons arranged into $N$ spin orbitals. The ground state wavefunction in Eq. (8.1) may be represented as a linear combinations of these arrangements,

$$|\Psi\rangle = \sum_{i=1}^{M} a_i |i\rangle . \tag{8.31}$$

The energies may be solved for variationally,

$$E = \min_{\{a_i\}} \frac{\langle\Psi| H |\Psi\rangle}{\langle\Psi|\Psi\rangle} \quad \rightarrow \quad H^{\mathrm{CI}} |\Psi\rangle = E |\Psi\rangle \tag{8.32}$$

where $H_{ij}^{\mathrm{CI}} = \langle i| H |j\rangle$. In chemistry this method is known as full configuration interaction (FCI).

FCI is strongly believed to be classically intractable because $M$ scales combinatorially with $N$ and $n$. Accordingly, a common classical approach is to truncate the expansion in Eq. (8.31) to include only configurations that represent a fixed number of excitations from a reference configuration. Though this work and recent work [195] discuss using different orbital basis choices, usually the reference is taken to be the Hartree-Fock state (this orbital basis is known in chemistry literature as the "canonical basis"). This approach defines a hierarchy of methods referred to as truncated configuration interaction (CI) which approach exactness as the number of excitations is increased to the FCI space spanned by $N-n$ excitations. Fixing the maximum number of excitations at $k$, combinatorics suggests that the number of basis functions in truncated CI scales as $\Theta\left(\binom{N-n}{k}\binom{n}{k}\right)$. Truncation to the level of single and double excitations is referred to as configuration interaction singles, doubles (CISD) and is used for several purposes in this paper. Finally, we note that the accuracy of truncated CI is extremely sensitive to the quality of the reference state and it is therefore difficult to determine when these methods are expected to approximate the ground state energy within even a fixed multiplicative error.

Since the error operator can be efficiently computed and normal-ordered in second quantized form, we suggest evaluating the expectation value of this operator on a classical ansatz for the ground state. In particular, we focus on the use of the configuration interaction ansatz. Figure 8.9 illustrates the utility of this idea by showing the discrepancy between actual error and the error from evaluation of the error operator using a classical ansatz. Figure 8.10 shows the extent to which the effective error is reduced using a classical ansatz.

Apart from estimating errors, CISD states may also be of use in coalescing schemes [224] which use the Hartree-Fock approximation to determine whether a term in the Hamiltonian

can be executed less frequently without significantly impacting the quality of the simulation. This process can substantially reduce the costs of simulating molecules with many small, but non–negligible, $h_{pqrs}$ terms but may fail if the Hartree-Fock approximation breaks down. In such cases, the use of CISD states may lead to superior coalescing schemes at the price of requiring more classical computing time to find the coalescing schedule.

Though the Hartree-Fock ansatz is usually not accurate enough to reduce error by an order of magnitude, the use of a truncated CI ansatz often exhibits enough accuracy to very substantially reduce effective error. While we focus on the CI ansatz to provide proof-of-principle, we believe that more intelligent truncation schemes can substantially increase ansatz accuracy without additional computational cost. For instance, the use of multi-reference methods has been shown to greatly improve the quality of the classical solution in many cases, especially near molecular dissociation limits where the exact electronic states become nearly degenerate [145].

The idea of using a classical ansatz to reduce the effective error in a quantum calculation is useful for two reasons. The first reason is that the error in a quantum simulation can usually be reduced by approximating the error with a classical ansatz at the CISD level of theory or greater, as demonstrated in Figure 8.10. The second (and perhaps more important) reason this technique is useful is that it gives a realistic a priori estimate of the error to expect in the quantum simulation (expected to be correct to at least an order of magnitude) which provides a methodology for selecting the number of Trotter steps required to obtain a desired precision. Finally, we point out that while the error operator might be computationally costly to compute (albeit, efficient in the polynomial-time scaling sense), Monte Carlo methods could be used to tractably sample the error operator expectation values with a classical ansatz.

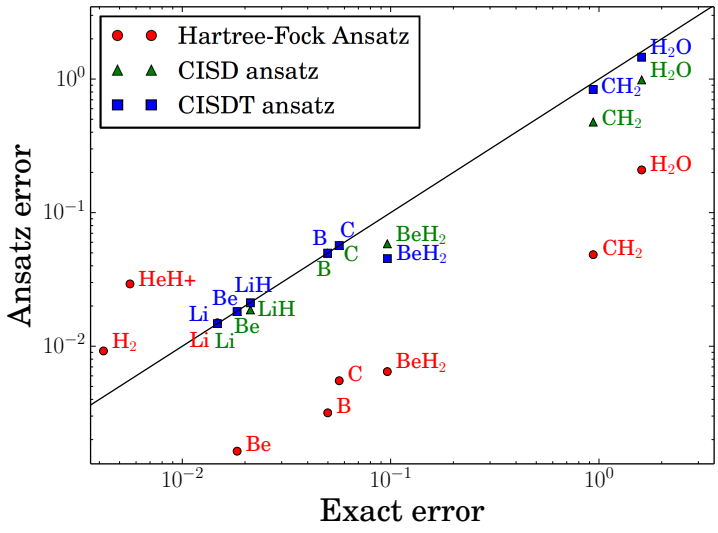



Figure 8.9: Magnitude of the Trotter error in the exact ground state against the magnitude of the error induced on a classical ansatz for the ground state. Truncated CI computations are only performed when inexact; e.g., we have not computed HF using CISD because the calculation is exact in STO-6G.

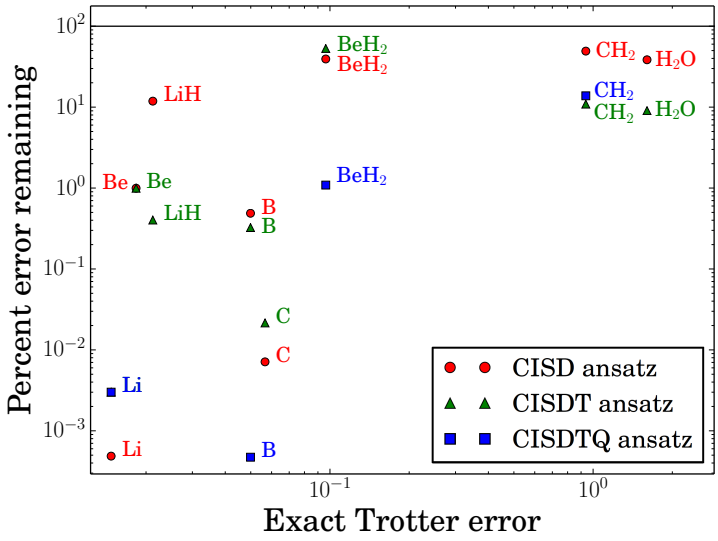



Figure 8.10: Magnitude of the Trotter error induced in the exact ground state against percentage of the error that remains after subtracting the ansatz error from the exact error. This plot is intended to indicate the reduction in effective error when using a classical ansatz estimate. A black line is drawn at one-hundred percent remaining. In all benchmarks, using these ansatzs reduces effective error. Note that the quadruple calculation is so accurate for Be and LiH that the effective error appears to be exact to within double precision.

### 8.3.1 Circuit for state preparation based on CI ansatz

In contrast to the Hartree-Fock states, CISD states are not computational basis states. Instead they are a linear combination of quantum states that are formed by single and double excitations away from a reference state which is often taken to be the Hartree-Fock state. Although the CISD state can be efficiently computed for a given electronic structure problem, preparing the state on a quantum computer is non-trivial. Here we present a method based on state-of-the-art multi-qubit synthesis methods to prepare the CISD state. Previous work has considered preparing this state using single qubit rotations and CNOT gates [217, 276, 296]. Such gate sets are unrealistic for fault tolerant quantum computing so we discuss the problem of compiling the state preparation circuit into Clifford and $T$ gates. In the following analysis we will take the cost of the circuit to be given by the number of $T$ gates because these gates are the most expensive gates to implement fault tolerantly in error correcting codes such as the surface code.

Let us begin by assuming the initial state for the quantum simulation (i.e. the state we wish to prepare) is of the form

$$|\psi\rangle = \sum_{k=1}^{D} \alpha_k |j_k\rangle , \tag{8.33}$$

where $j_k$ is a sequence of computational basis vectors that spans the space that state has support over and $D$ is the dimension of that space.

It is unrealistic to assume that the state $|\psi\rangle$ will be exactly preparable using gates from the Clifford + $T$ gate library. Instead, the initial state will typically have to be approximated using these circuit elements. For years the Solovay-Kitaev algorithm provided the best known method for solving this approximation problem, but recently more advanced methods based on number theoretic results have provided much more efficient ways of performing this decomposition [170, 234, 41].

Therefore the problem of finding the best sequence of Clifford and $T$ gates to approximate a multi-qubit unitary reduces to the following problem

1. Find integers $x_0, x_1, y_0, y_1$ such that

$$U_{p,q} \approx \tilde{U}_{p,q} = \frac{x_0 + x_1\sqrt{2} + iy_0 + iy_1\sqrt{2}}{\sqrt{2}^m}.$$

   and $\tilde{U}$ is a unitary that can be exactly synthesized using elements from the gate library.

2. Find a sequence of Clifford and $T$ gates that exactly implements $\tilde{U}$.

Note that because we are interested in preparing a state, not implementing a multi-qubit unitary, only the first column of $U$ needs to be approximated. In particular, the first column of $\tilde{U}$ should approximate $|\psi\rangle$ to within a fixed error tolerance $\delta$.

Before proceeding it is necessary to briefly review number theoretic approaches to multi-qubit circuit synthesis using Clifford and $T$ gates. The key insight behind this strategy is that the unitary matrices that can be prepared with such circuits take on a very special form. The form can easily be seen from the Hadamard and $T$ gates,

$$H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \qquad T = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1+i}{\sqrt{2}} \end{bmatrix}. \tag{8.34}$$

It is then clear that any unitary matrix formed by a sequence of $H$ and $T$ gates will consist of matrix elements that are of the form

$$\tilde{U}_{i,j} = \frac{x_0 + x_1\sqrt{2} + iy_0 + iy_1\sqrt{2}}{\sqrt{2}^m}, \tag{8.35}$$

for integer $x_0, x_1, y_0, y_1$. Since the remainder of the gate set consists of CNOT gates and Pauli gates which have (complex) integer valued matrix elements, it is then clear that every

unitary that can be formed by the gate library also has matrix elements whose denominators are powers of $\sqrt{2}$ and whose numerators are in the ring of Gaussian integers $\mathbb{Z}[1/\sqrt{2}, i]$.

Just like ordinary fractions, these fractions also can be reduced. This notion of reducing a fraction manifests itself as the least denominator exponent $k$. In order to understand this concept concretely, it is necessary to introduce some terminology. Let $\omega = e^{i\pi/4}$ and

$$\mathbb{Z}[\omega] = \{a\omega^3 + b\omega^2 + c\omega + d | a, b, c, d \in \mathbb{Z}\}. \tag{8.36}$$

Similarly, if we let $\mathbb{D} = \{a2^{-b} | a, b \in \mathbb{Z}\}$ denote the ring of dyadic fractions then we can express the ring $\mathbb{Z}[1/\sqrt{2}, i]$ as

$$\mathbb{D}[\omega] = \{a\omega^3 + b\omega^2 + c\omega + d | a, b, c, d \in \mathbb{D}\}. \tag{8.37}$$

Then for every $t \in \mathbb{D}[\omega]$ there is a notion of a *least denominator exponent* that describes the fraction in Eq. (8.35) and uses the smallest value of $m$ possible while requiring that $x_0, x_1, y_0, y_1$ are integer. Or more formally, the least denominator exponent, $k$, is the smallest non-negative integer such that $t\sqrt{2}^k \in \mathbb{Z}[\omega]$.

The smallest denominator exponent measures the precision in the approximation $U \approx \tilde{U}$ because Eq. (8.35) allows arbitrary complex numbers to be represented with zero error in the limit as $k \to \infty$. This means that the value of $k$ used in the rounding process of the first column of $U$ is a key property for characterizing the complexity of the state preparation. In fact, the problem of bounding the error in this approximation problem as a function of $k$ has already been solved by Kliuchnikov [170]:

$$\| (U - \tilde{U}) |0\rangle \| \leq 2(D+2)2^{-4k} + 2\sqrt{2(D+2)}2^{-2k}, \tag{8.38}$$

where $D$ is the number of nonzero components of the state $|\psi\rangle = U|0\rangle$. As a technical point, the dimension of $\tilde{U}$ is at most $D+2$ rather than $D$ because the first column of $U$

must have *at least two zero-valued components* in order to guarantee that a solution exists to the Diophantine equation for $\tilde{U}$. This requires enlarging the Hilbert space dimension by two in the worst case scenario, which may require adding at most an additional qubit. However, the CISD state vector will likely have many zero valued components so this extra qubit will often not be needed in practice.

Using Eq. (8.38) we see that the state preparation error can be made less than $\delta$ by choosing

$$k = \left\lceil \frac{1}{4} \left[ 1 + \log_2 \left( \frac{D+2}{(\sqrt{1+\delta}-1)^2} \right) \right] \right\rceil . \tag{8.39}$$

This means that if $D$ is polynomial in $n$ then $k \in O(\log(n/\delta))$.

Once the unitary $\tilde{U}$ has been found then the task of decomposing the unitary into fundamental operations remains a non-trivial problem. This problem is addressed by Giles and Selinger in [119]. The idea behind this approach is to decompose $\tilde{U}$ into a series of two level unitary operations. These two level unitary operations are then implemented using a Clifford circuit and a series of controlled operations to map each two level subspace to a single qubit. This process involves first identifying pairs of levels that can be simplified and then performing circuits of the form $H^w T^x H^y T^z$ to the two level subspace such that the denominator exponent is systematically reduced. Once the least denominator exponent is reduced to 0 then the subspace either takes the form $[\omega^p, 0]^T$ or $[0, \omega^p]^T$ for integer $p$. Thus, the inverse of the state preparation circuit can be found (up to a global phase) by performing this reduction process iteratively of the $D$ dimensional initial state until only one nonzero component remains and then mapping this component to $|0\rangle$ using a Clifford circuit and a multiply controlled not gate.

At most $k$ reduction steps are needed to reduce each two level subspace and there are at most $(D+2)-1$ subspaces that must be looped through. Therefore, there are at most

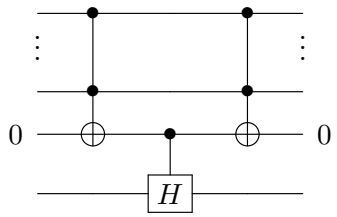
$k(D+1)$ reduction steps taken. Each reduction step consists of applying at most two $H$ gates and two $T^x$ gates to each subspace, as well as a multiply controlled not gate to map the final state to one proportional to $|0\rangle$. Hence, in order to assess the cost of the algorithm we need to compute the costs of each of these gates.

Let us imagine that we need to perform a gate on the subspace $\text{span}(|j\rangle, |k\rangle)$. We want to map this to $\text{span}(|2^n-1\rangle, |2^n-2\rangle)$ so that the gate can be applied to the last qubit. By performing a sequence of $\mathcal{O}(n)$ $X$ gates, we can map

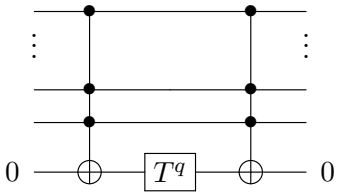
$$\text{span}(|j\rangle, |k\rangle) \to \text{span}(|j \oplus k \oplus 2^n - 1\rangle, |2^n - 1\rangle),$$

where $\oplus$ is bitwise exclusive or. There are two cases that we need to consider. If $j \oplus k = 1 \mod 2$ then the least significant bit of $j \oplus k \oplus 2^n - 1$ is 0. This means that the state $|j \oplus k \oplus 2^n - 1\rangle$ can be mapped to $|2^n - 2\rangle$ using a sequence of $n-1$ zero-controlled not gates while not affecting $|2^n - 1\rangle$. Otherwise, if $j \oplus k = 1 \mod 2$ then we can reduce this case by finding the least significant bit where $j$ and $k$ differ and swap that bit with the least significant bit. Since $|2^n - 1\rangle$ is an eigenstate of the swap operator, the swap does not affect that vector. Hence in either case we can perform the subspace mapping using $\mathcal{O}(n)$ Clifford operations.

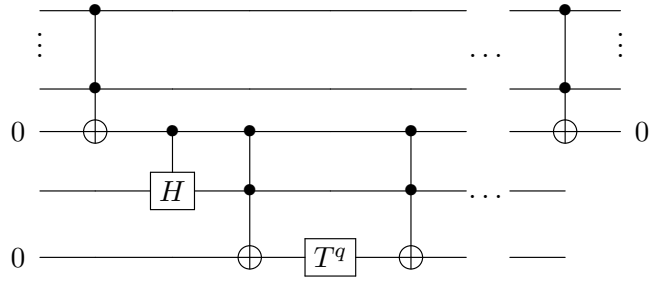
In order to apply the $H$ and $T$ gates required by the synthesis algorithm on the correct qubits, we need to implement controlled variants of these circuits. There are many constructions for these controlled gates [24, 155, 244]. Here we anticipate that the cost of state preparation for the CISD state will be sub-dominant to the cost of the simulation. This means that minimizing the number of qubits needed is an important design goal. Let us define $\Lambda_m(G)$ to be the $m$-controlled version of the gate $G$. Then the gate $\Lambda_m(H)$ can be implemented using two $\Lambda_{n-1}(X)$ gates, a $\Lambda_1(H)$ gate and an ancilla qubit,




Controlled $T^q$ gates can be performed similarly,




The resulting circuits can be further optimized by noting that many of the Toffoli gates needed to perform the reductions of the least denominator exponent are redundant. In particular we can express the simplified reduction circuit as,




The gate $\Lambda_1(H)$ requires 2 $T$-gates [119], and the Toffoli gates can be implemented, up to an irrelevant phase, using 4 $T$-gates [155, 244] and an ancilla qubit. The entire process requires at most $N + 4$ qubits, which is typically less memory than is required for the quantum simulation and eigenvalue estimation phases of the algorithm. This means that the additional four qubits required for the state preparation algorithm will not impact the memory requirements of the overall simulation algorithm.

For the present problem, the CISD state is in $\mathbb{C}^{2^{N+1}}$ (recall one additional qubit is needed to ensure a solution to the norm equations for synthesis). This means that we also

need to consider the cost of implementing $\Lambda_N(X)$ gates. Although highly time-efficient constructions for the multiply controlled circuits can be made using the circuits of [155], they require a large number of qubits. In order to ensure that the space complexity of state preparation does not dominate the algorithm, we use the less time-efficient construction of Barenco et al [24] to compile the $\Lambda_N(X)$ gates. Using Corollary 7.4 from [24] and the $\Lambda_2(iX)$ gate from [155, 244] to implement the Toffoli gate, the cost of implementing such circuits is at most

$$T_{\text{count}}\big(\Lambda_N(iX)\big) \leq 32(N-3). \tag{8.40}$$

At most $N+4$ qubits, where $N \geq 5$, are needed to implement these gates [24].

The reduction of each of the two dimensional subspaces requires two steps. First, the application of the $\Lambda_N(X)$ gates to mark the subspace and a sequence of $k$ controlled operations to reduce the denominator exponent of that subspace. This reduction process requires $k$ steps, each of which involves at most two $\Lambda_1(H)$ and two $\Lambda_2(T^q)$. Including the cost of the two $\Lambda_N(X)$ gates, the total cost of the reduction is at most $22k + 64(N-3)$ $T$-gates. At most $D+1$ reduction steps are required in this process as well as potnentially a swap of the final state into the $|0^{N+1}\rangle$ state (which can be performed using Clifford operations). Thus the overall $T$–count for this process is

$$(22k + 64(N-3))(D+1). \tag{8.41}$$

This also is the $T$–count for for preparing the CISD state from the $|0^{N+1}\rangle$ state because the necessary circuit can be found by taking the Hermitian conjugate of the resultant gate sequence. Thus using Eq. (8.39) the total number of non-Clifford operations required in the state preparation scales at most as

$$\mathcal{O}\left(D\log\left(\frac{D}{\delta}\right)+ND\right) \tag{8.42}$$

If the approach of Wang et al [276], coupled with recent methods for decomposing single qubit rotations into $T$ gates, is used to prepare the CISD state then the resultant $T$-count scales at most as $\tilde{\mathcal{O}}(2^{n_e}N^{n_e}/n_e!)\log(1/\delta)$. If $n_e \approx N/2$ then this method is inefficient, whereas ours is not since $D \in \mathcal{O}(N^4)$ for CISD states. If $n_e \leq 3$ then the method of Wang et al does provide superior scaling as $N$ increases, though cases where $n_e \leq 3$ and $N$ is large may be rare. In contrast, the method of Ortiz et al [217] requires $\tilde{\mathcal{O}}(D^2N^2\log(1/\delta)$ gates, which is nearly quadratically slower than our method.

As a final point, the cost of the state preparation algorithm is $\mathcal{O}(N^5)$ in worst case scenarios. This can be comparable to, or greater than, the cost of quantum simulation in the limit of large $N$. This means that using a naïve CISD approximation in cases with half filling may seriously degrade the performance of the algorithm. This means that in order to see the performance advantages promised by recent algorithms, which have scaling near $\mathcal{O}(N^4)$, sophisticated state preparation methods are needed in cases where the Hartree-Fock state has poor overlap with the FCI ground state.

## 8.4 Conclusion

Our work calls into question the basic assumption that the error in Trotter-Suzuki based methods for simulating quantum chemistry is *explicitly* a function of the number of spin orbitals used to represent the system. We find through numerical evidence that such errors do not seem to be directly related to the number of spin orbitals in the system for small molecules. We observe this lack of correlation for a variety of orbital bases including local,

canonical and natural orbitals. Instead, we see that chemical features such as the maximum nuclear charge is a strong indicator of the complexity of a simulation. We argue that the errors should scale as $\mathcal{O}(Z_{\max}^6)$ for an atomic orbital basis, which is in close agreement with the scaling observed numerically. We also observe that some atoms, such as Oxygen, Fluorine and Neon, have vanishingly small Trotter errors despite available error bounds predicting large Trotter errors for these molecules. We show that this discrepancy can be understood as a consequence of the large filling fraction for these molecules. This suggests that chemical features of a molecule may be much better predictors of the number of Trotter steps needed in a simulation than the number of spin orbitals assigned to the molecule.

We further analyze the errors and see that the discrepancy between the observed Trotter error and the norm of the error operator does not arise from random cancellation. Indeed, the errors observed are much greater than what would be expected if the ground state were a Haar random state that was chosen independently from the eigenvectors of the error operator. Furthermore, we observe that the distribution of eigenvalues of the error operator is highly structured and has many near-zero eigenvalues, which likely is the cause of the orders of magnitude separation between the Trotter error and the norm of the error operator.

We also use the error operator to improve quantum simulation methods by providing a computationally efficient algorithm for estimating the error in a simulation. This leads to two applications: ($a$) compensating for Trotter error in a quantum simulation by subtracting the prediction off the result and ($b$) predicting the number of Trotter steps needed in a simulation. Finally, we provide a quantum algorithm for preparing CISD states that is polynomially more efficient than existing methods and may provide a viable alternative to adiabatic state preparation in cases where the Hartree-Fock approximation to the ground

state leads to poor success probability.

There are several natural avenues of inquiry that this work reveals. First, although this work shows strong numerical evidence for small molecules we do not have sufficient evidence to state conclusively that the error in the Trotter-Suzuki formula is independent of $N$ in the asymptotic limit. Larger numerical experiments may be needed to shed more light on the scaling of Trotter–Suzuki errors in this regime. Secondly, Ferredoxin is often suggested as a strong candidate for quantum chemistry simulation but $Fe_2S_2$ has large nuclear charges which make it a challenging molecule from the perspective of simulation. This suggests that there may be other large organic molecules with smaller nuclear charges that may be even more natural targets for quantum simulation. Finally, although our work has suggested that the number of spin orbitals in a molecule may not uniquely characterize the cost of a quantum chemistry simulation, it does not provide a simple criteria for determining which molecules are easy or hard to simulate. Finding molecular features, beyond the maximum nuclear charge and the filling fraction, that can be used to predict the relative difficulty of simulation would not only constitute an important step forward for quantum chemistry simulations but would also be an important contribution to quantum chemistry as a whole.

## Acknowledgements

The authors thank Matthew Hastings and David Gosset for helpful conceptual discussions, Peter Love for comments on an early manuscript, John Parkhill for conversations regarding classical electronic structure methods and Vadym Kliuchnikov for discussions about preparation of the CISD state. J.M. is supported by the Department of Energy Computational Science Graduate Fellowship under grant DE-FG02-97ER25308. A.A-G. appreciates the support of the Air Force Office of Scientific Research under award FA9550-12-1-0046.

## 8.5 Appendix

### 8.5.1 Computation of Haar Expectations

In order to determine whether the error cancellations observed for ground state quantum simulations arise because of properties of the eigenstates of the Hamiltonian, we need to determine whether these results would also be typical of random vectors. Here we provide a derivation, for completeness, of the Haar expectation value and variance of the $|\langle v|k\rangle|^2$.

In the following we will take $k$ to be fixed and $v$ to represent the Haar random variable. We will also use the convention that $\mathbb{E}_H$ denotes the expectation value of a quantity over a set of Haar random vectors, and $\mathbb{V}_H$ denotes the variance over the set. To be clear, $\mathbb{E}_H|\langle v|k\rangle|^2 = \int_{U\in\text{Haar}} |\langle 0| U^\dagger |k\rangle|^2 \mathrm{d}U$.

We wish to compute the variance,

$$\mathbb{V}_H(|\langle v|k\rangle|^2) = \mathbb{E}_H(|\langle v|k\rangle|^4) - \mathbb{E}_H(|\langle v|k\rangle|^2)^2, \tag{8.43}$$

of the square of the overlap of an arbitrary Haar random vector, $|0\rangle$, with an eigenvector of an arbitrary Hermitian operator (in this case, the Trotter error operator), $|k\rangle$. We begin by stating the correspondence,

$$|v\rangle\langle v| = U|0\rangle\langle 0|U^\dagger \tag{8.44}$$

where $U$ is the unitary Gram matrix which affects a basis transformation into the error operator eigenbasis (for instance), and $|v\rangle$ represents $|0\rangle$ in the error operator eigenbasis. We are interested in the projection of this state onto an eigenvector of the error operator,

$$a_k = \langle k|v\rangle = \langle k|U|0\rangle \tag{8.45}$$

$$\begin{aligned} |a_k|^2 &= \langle k|U|0\rangle\langle 0|U^\dagger|k\rangle \\ &= \mathrm{tr}\left[|k\rangle\langle k|U|0\rangle\langle 0|U^\dagger\right]. \end{aligned} \tag{8.46}$$

We compute this trace in two steps. From the unitary invariance of the Haar measure we have that

$$\int_{U(n)} \left[ U \left| 0 \right\rangle \left\langle 0 \right| U^\dagger \right] \mathrm{d}U = \frac{I}{2^N} \tag{8.47}$$

Therefore

$$\mathbb{E}_H(|a_k|^2) = \mathrm{tr}\left[ |k\rangle\langle k| \frac{I}{2^n} \right] = \frac{1}{2^N}. \tag{8.48}$$

Thus, $\mathbb{E}_H(|a_k|^2) = \frac{1}{2^N}$, and hence $\mathbb{E}_H(|a_k|^2)^2 = \frac{1}{2^{2N}}$.

Focusing on the remaining component of the variance,

$$\begin{aligned} |a_k|^4 &= \langle k| U |0\rangle \langle 0| U^\dagger |k\rangle \langle k| U |0\rangle \langle 0| U^\dagger |k\rangle \\ &= \mathrm{tr}\left[ \left( |k\rangle \langle k| \right)^{\otimes 2} U^{\otimes 2} \left( |0\rangle \langle 0| \right)^{\otimes 2} U^{\dagger \otimes 2} \right]. \end{aligned} \tag{8.49}$$

To further evaluate the trace we follow the treatment in [142] which uses the spectral theorem to derive orthogonal projectors onto symmetric and antisymmetric subspaces. This begins by defining a flip operator, $\mathbb{F} \in \mathbb{C}^{2^{2N} \times 2^{2N}}$,

$$\mathbb{F}\left( |\psi\rangle \otimes |\varphi\rangle \right) = |\varphi\rangle \otimes |\psi\rangle . \tag{8.50}$$

From this definition it is clear that

$$\mathbb{F} = \pi_{\mathrm{sym}} - \pi_{\mathrm{antisym}} \tag{8.51}$$

$$I^{\otimes 2} = \pi_{\mathrm{sym}} + \pi_{\mathrm{antisym}}. \tag{8.52}$$

Thus,

$$\pi_{\mathrm{sym}} = \frac{1}{2}\left( I^{\otimes 2} + \mathbb{F} \right) \tag{8.53}$$

$$\pi_{\mathrm{antisym}} = \frac{1}{2}\left( I^{\otimes 2} - \mathbb{F} \right). \tag{8.54}$$

Since $\text{tr}\left[I^{\otimes 2}\right] = 2^{2N}$ and $\text{tr}(\mathbb{F}) = 2^N$,

$$\text{tr}\left[\pi_{\text{sym}}\right] = \frac{2^N\left(2^N+1\right)}{2} \tag{8.55}$$

$$\text{tr}\left[\pi_{\text{sym}}\right] = \frac{2^N\left(2^N-1\right)}{2}. \tag{8.56}$$

Since $|0\rangle\langle 0|\otimes|0\rangle\langle 0|$ is entirely symmetric, it is straight forward to see from unitary invariance that

$$\int_{U(n)}\left[U^{\otimes 2}\left(|0\rangle\langle 0|\right)^{\otimes 2}U^{\dagger\otimes 2}\right]\mathrm{d}U = \frac{2}{2^N\left(2^N+1\right)}\pi_{\text{sym}}. \tag{8.57}$$

Thus,

$$\begin{aligned}\mathbb{E}_H(|a_k|^4) &= \frac{2}{2^N\left(2^N+1\right)}\text{tr}\left[\pi_{\text{sym}}\left(|k\rangle\langle k|\right)^{\otimes 2}\right] \\ &= \frac{2}{2^N\left(2^N+1\right)}.\end{aligned} \tag{8.58}$$

Finally, we arrive at the variance of $|a_k|^2$,

$$\mathbb{V}_H(|a_k|^2) = \frac{2}{2^N\left(2^N+1\right)} - \frac{1}{2^{2N}}. \tag{8.59}$$

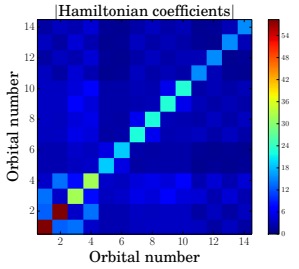
This gives us the variance in the $|a_k|^2$ terms, which in turn allows us to find the deviation from the expected error for a quantum chemistry simulation. The key point here is that the standard deviation is on the order of the expectation value which means that we expect relatively large fluctuations in the probabilities that correspond to particular eigenvalues of the Trotter error operator. Hence we do not expect a concentration of measure result to hold in high dimensional spaces.

### 8.5.2 Contributions of orbitals to Trotter error operator

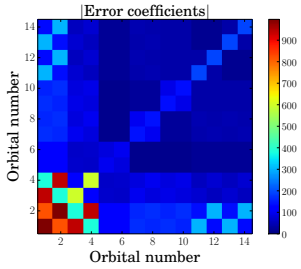
In Section 8.2 we provided evidence that transitions involving the inner-most electrons contribute most to the error in the quantum simulation. This is perhaps surprising given that transitions involving the valence electrons, rather than the core electrons, are typically more relevant for understanding the properties of a molecule. We present additional numerical results in Figure 8.11 that examine this for water and Beryllium hydride. These results confirm the intuition developed earlier that interactions involving the two inner–most orbitals significantly impact the errors in the ground state energy.

We also observe a rough correlations between the magnitudes of the error coefficients in the expansion of the error operator and their contribution to the ground state error in the local orbital basis. This suggests that looking at the contribution of the inner orbitals is the most significant for the error for these molecules and that the aggregate contributions of the error coefficients correlates roughly with the ground state error.

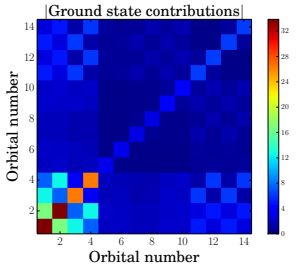
The analogous data for natural orbital basis in Figure 8.11 defies this approximate correspondence for both water and Beryllium hydride. For the case of water, interactions that involve orbitals 7 and 8 are the second-largest contributors to the error in the ground state energy. The significance of these transitions is not apparent in the corresponding plots of the magnitude of the Hamiltonian coefficients nor the magnitudes of the error coefficients in the expansion of the error operator. Similarly, for Berylium Hydride, interactions involving orbitals 7 and 8 may be expected to have a significant impact on the error in the ground-state energy but the data suggests that they do not. These results underscore the challenges faced when attempting to understand the nature of the error operator from solely looking at the magnitudes of the error coefficients.



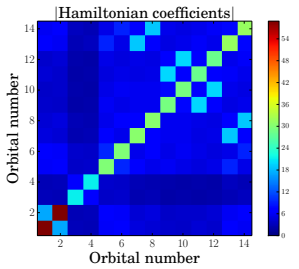

(a) Water, local basis, Hamiltonian coefficients.



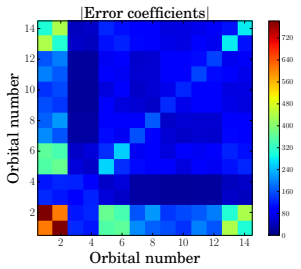

(b) Water, local basis, error coefficients.



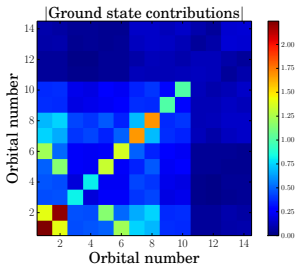

(c) Water, local basis, error contributions.



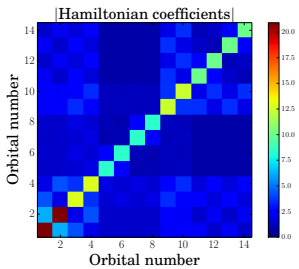

(d) Water, natural basis, Hamiltonian coefficients.



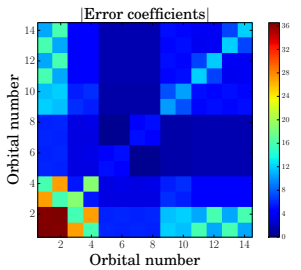

(e) Water, natural basis, error coefficients.



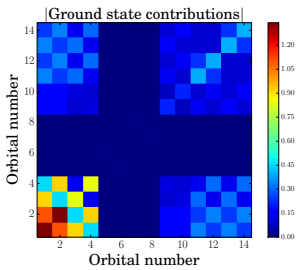

(f) Water, natural basis, error contributions.



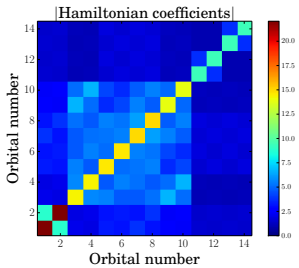

(g) Beryllium hydride, local basis, Hamiltonian coefficients.



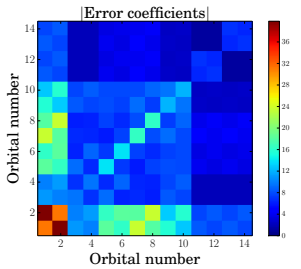

(h) Beryllium hydride, local basis, error coefficients.



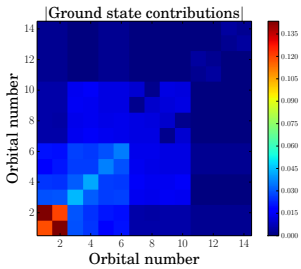

(i) Beryllium hydride, local basis, error contributions.




(j) Beryllium hydride, natural basis, Hamiltonian coefficients.




(k) Beryllium hydride, natural basis, error coefficients.




(l) Beryllium hydride, natural basis, error contributions.

Figure 8.11: These plots show the coefficients of normal-ordered terms in the Hamiltonian and error operator as well as expectation values of the error operator terms for the ground state. The terms are binned according to the orbitals involved in the term. This plot shows the marginal distribution of the magnitudes of these terms.

# Chapter 9

# Exploiting Locality in Quantum Chemistry Simulation

Apart from minor modifications, this chapter originally appeared as [195]:

"Exploiting Locality in Quantum Computation for Quantum Chemistry". Jarrod McClean, Ryan Babbush, Peter Love and Alán Aspuru-Guzik. *Journal of Physical Chemistry Letters*. Volume 5, Number 24: 4368-4380. 2014.

## Abstract

Accurate prediction of chemical and material properties from first principles quantum chemistry is a challenging task on traditional computers. Recent developments in quantum computation offer a route towards highly accurate solutions with polynomial cost, however this solution still carries a large overhead. In this perspective, we aim to bring together known results about the locality of physical interactions from quantum chemistry with ideas from quantum computation. We show that the utilization of spatial locality combined with the Bravyi-Kitaev transformation offers an improvement in the scaling of known quantum algorithms for quantum chemistry and provide numerical examples to help illustrate this point. We combine these developments to improve the outlook for the future of quantum chemistry on quantum computers.

## 9.1 Introduction

Within chemistry, the Schrödinger equation encodes all information required to predict chemical properties ranging from reactivity in catalysis to light absorption in photovoltaics. Unfortunately the exact solution of the Schrödinger equation is thought to require exponential resources on a classical computer, due to the exponential growth of the dimensionality of the Hilbert space as a function of molecular size. This makes exact methods intractable for more than a few heavy atoms [262].

Richard Feynman first suggested that this scaling problem might be overcome if a more natural approach was taken [107]. Specifically, instead of painstakingly encoding quantum information into a classical computer, one may be able to use a quantum system to naturally represent another quantum system and bypass the seemingly insurmountable storage requirements. This idea eventually developed into the field of quantum computation, which is now believed to hold promise for the solution of problems ranging from factoring numbers [251] to image recognition [209, 13] and protein folding [221, 18].

Initial studies by Aspuru-Guzik et. al. showed that these approaches might be particularly promising for quantum chemistry [11]. There have been many developments both in theory [77, 243, 296] and experimental realization [175, 12, 223, 279] of quantum chemistry on quantum computers. The original gate construction for quantum chemistry introduced by Whitfield et al. [285] was recently challenged as too expensive by Wecker et al. [281]. The pessimistic assessment was due mostly to the extrapolation of the Trotter error for artificial rather than realistic molecular systems, as was analyzed in detail in a followup study by many of the same authors [224]. They subsequently improved the scaling by means of several circuit enhancements [140]. The analysis of the Trotter error on realistic molecules in combination with their improvements led to a recent study where an estimate

of the calculation time of $Fe_2S_2$ was reduced by orders of magnitude [224]. In this paper, we further reduce the scaling by exploiting the locality of physical interactions with local basis sets as has been done routinely now in quantum chemistry for two decades [64, 121]. These improvements in combination with others make quantum chemistry on a quantum computer a very attractive application for early quantum devices. We describe the scaling under two prominent measurement strategies, quantum phase estimation and Hamiltonian averaging, which is a simple subroutine of the recently introduced Variational Quantum Eigensolver approach [223].

Additionally, recent progress in accurate and scalable solutions of the Schrödinger equation on classical computers has also been significant [64, 10, 121, 145, 247, 52]. Some of these results have already appeared in the quantum computation literature in the context of in depth studies of state preparation [277, 272]. A general review of quantum simulation [59, 116] and one on quantum computation for chemistry [163] cover these topics in more depth. However many developments that utilize fundamental physical properties of the systems being studied to enable scalability have not yet been exploited.

In this study, we hope to bring to light results from quantum chemistry as well as their scalable implementation on quantum computers. We begin by reviewing the standard electronic structure problem. Results based on the locality of physical interactions from linear scaling methods in quantum chemistry are then introduced with numerical studies to provide quantification of these effects. A discussion of the resulting impact on the most common quantum algorithms for quantum chemistry follows. We also investigate instances where a perfect oracle is not available to provide input states, demonstrating the need for advances in state preparation technology. Finally, we conclude with an outlook for the future of quantum chemistry on quantum computers.

### 9.1.1 Electronic structure problem

To frame the problem and set the notation, we first briefly introduce the electronic structure problem of quantum chemistry [145]. Given a set of nuclei with associated charges $\{Z_i\}$ and a total charge (determining the number of electrons), the physical states of the system can be completely characterized by the eigenstates $\{|\Psi_i\rangle\}$ and corresponding eigenvalues (energies) $\{E_i\}$ of the Hamiltonian $H$

$$H = -\sum_i \frac{\nabla_{R_i}^2}{2M_i} - \sum_i \frac{\nabla_{r_i}^2}{2} - \sum_{i,j} \frac{Z_i}{|R_i - r_j|} + \sum_{i,j>i} \frac{Z_i Z_j}{|R_i - R_j|} + \sum_{i,j>i} \frac{1}{|r_i - r_j|} \tag{9.1}$$

where we have used atomic units, $\{R_i\}$ denote nuclear coordinates, $\{r_i\}$ electronic coordinates, $\{Z_i\}$ nuclear charges, and $\{M_i\}$ nuclear masses. Owing to the large difference in masses between the electrons and nuclei, typically the Born-Oppenheimer approximation is used to mitigate computational cost and the nuclei are treated as stationary, classical point charges with fixed positions $\{R_i\}$. Within this framework, the parametric dependence of the eigenvalues on $\{R_i\}$, denoted by $\{E(\{R_i\})_j\}$ determines almost all chemical properties, such as bond strengths, reactivity, vibrational frequencies, etc. Work has been done in the determination of these physical properties directly on a quantum computer [162].

Due to the large energy gaps between electronic levels with respect to the thermal energy scale $k_B T$, it typically suffices to study a small subset of the eigenstates corresponding to the lowest energies. Moreover, for this reason, in many molecules the lowest energy eigenstate $|\Psi_0\rangle$, or ground state, is of primary importance, and for that reason it is the focus of many methods, including those discussed here.

### 9.1.2 Second quantized Hamiltonian

Direct computation in a positional basis accounting for anti-symmetry in the wavefunction while using the Hamiltonian described is referred to as a first quantization approach

and has been explored in the context of quantum computation [162, 280, 283]. The first quantized approach has also been realized in experiment [187]. One may also perform first quantized calculations in a basis of slater determinants. This was introduced as a representation of the electronic wavefunction by qubits in [11] (the compact mapping) and the efficiency of time evolution in this basis was recently shown [264]. The second quantized approach places the antisymmetry requirements on the operators. After choosing some orthogonal spin-orbital basis $\{\varphi_i\}$ with a number of terms $M$, the second quantized Hamiltonian may be written as [145]

$$H = \sum_{pq} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s \tag{9.2}$$

with coefficients determined by

$$h_{pq} = \int d\sigma \ \varphi_p^*(\sigma) \left( -\frac{\nabla_r^2}{2} - \sum_i \frac{Z_i}{|R_i - r|} \right) \varphi_q(\sigma) \tag{9.3}$$

$$h_{pqrs} = \int d\sigma_1 \ d\sigma_2 \ \frac{\varphi_p^*(\sigma_1)\varphi_q^*(\sigma_2)\varphi_r(\sigma_1)\varphi_s(\sigma_2)}{|r_1 - r_2|} \tag{9.4}$$

where $\sigma_i$ now contains the spatial and spin components of the electron, $\sigma_i = (r_i, s_i)$. The operators $a_p^\dagger$ and $a_r$ obey the fermionic anti-commutation relations

$$\{a_p^\dagger, a_r\} = \delta_{p,r} \tag{9.5}$$

$$\{a_p^\dagger, a_r^\dagger\} = \{a_p, a_r\} = 0 \tag{9.6}$$

For clarity, we note that the basis functions used in quantum chemistry (such as atom-centered Gaussians) are frequently parameterized on the nuclear coordinates $\{R_i\}$, which can result in a dependence on the nuclear positions of the electronic integral terms $\{h_{pqrs}\}$. For notational simplicity the dependence of the integrals on the nuclear positions in this work will remain implied.

### 9.1.3 Spatial locality

It is clear by inspection that the maximum number of terms in the second-quantized Hamiltonian scales as $O(M^4)$. $M$ can be quite large to reach chemical accuracy for systems of interest, and the number of terms present in the Hamiltonian is a dominant cost factor for almost all quantum computation algorithms for chemistry. However, due to the locality of physical interactions, one might imagine that many of the terms in the Hamiltonian are negligible relative to some finite precision $\epsilon$. While this depends on the basis, it is this observation that forms the foundation for the linear-scaling methods of electronic structure such as linear scaling density functional theory or quantum Monte Carlo [10, 290, 11]. That is, in a local basis, the number of non-negligible terms scales more like $O(M^2)$, and advanced techniques such as fast multipole methods techniques can evaluate their contribution in $O(M)$ time.

These scaling properties are common knowledge within the domain of traditional quantum chemistry, however they have not yet been exploited within the context of quantum computation. They are clearly vitally important for the correct estimate of the asymptotic scaling of any method [11, 285, 77, 281]. For that reason, we review the origin of that scaling here for the most common and readily available local basis, the Gaussian atomic orbital basis. We follow loosely the explanation presented by Helgaker, Jørgensen, and Olsen [145], and refer readers to this text for additional detail on the evaluation of molecular integrals in local basis sets. The two elements we will consider here are the cutoffs due to exponentially vanishing overlaps between Gaussians basis functions and a bound on the value of the largest integral.

By far the most common basis used in electronic structure calculations is a set of atom-centered Gaussian (either Cartesian or "Pure" spherical) functions. While the precise result

can depend on the angular momentum associated with the basis function, for simplicity, consider only Gaussian $S$ functions, which is defined by

$$|G_a\rangle = \exp\left(-ar_A^2\right) \tag{9.7}$$

where $r_A$ is the vector from a point $A$ which defines the center of the Gaussian. One property of Gaussian functions that turns out to be useful in the evaluation of molecular integrals is the Gaussian product rule. This rule states simply that the product of two spherical Gaussian functions may be written in terms of a single spherical Gaussian function on the line segment connecting the two centers. Consider two spherical Gaussian functions, $|G_a\rangle$ and $|G_b\rangle$ separated along the $x$-axis.

$$\exp\left(-ax_A^2\right)\exp\left(-bx_B^2\right) = K_{ab}^x \exp\left(-px_p^2\right) \tag{9.8}$$

where $K_{ab}^x$ is now a constant pre-exponential factor

$$K_{ab}^x = \exp\left(-\mu X_{AB}^2\right) \tag{9.9}$$

and the total exponent $p$, the reduced exponent $\mu$, and the Gaussian separation $X_{AB}$ are given by

$$p = a + b \tag{9.10}$$

$$\mu = \frac{ab}{a+b} \tag{9.11}$$

$$X_{AB} = A_x - B_x \tag{9.12}$$

That is, the product of two spherical Gaussians is a third Gaussian centered between the original two that decays faster than the original two functions, as given by the total exponent $p$. The overlap integral of two spherical Gaussian $S$ functions may be obtained through application of the Gaussian product rule after factorizing into the three Cartesian

dimensions, followed by Gaussian integration and is given by

$$S_{ab} = \langle G_a | G_b \rangle = \left( \frac{\pi}{a+b} \right)^{3/2} \exp\left( -\frac{ab}{a+b} R_{AB}^2 \right) \tag{9.13}$$

where $R_{AB}$ is the distance between the Gaussian centers $A$ and $B$. Clearly this integral decays exponentially with the square of the distance between centers, and one may determine a distance $d_s$ such that beyond that distance, the integrals will be smaller than $10^{-k}$ in magnitude.

$$d_s = \sqrt{a_{\min}^{-1} \log \left[ \left( \frac{\pi}{2a_{\min}} \right)^3 10^{2k} \right]} \tag{9.14}$$

where $a_{\min}$ is the minimal Gaussian exponent $a$ (most diffuse function) in the set of Gaussian basis functions $\{|G_a\rangle\}$. While the exact decay parameters will depend on the basis set, it is generally true from this line of reasoning that there is a characteristic distance, beyond which all overlap integrals are negligible. This means that the number of interactions per basis function becomes fixed, resulting in a linear number of significant overlap integrals. As kinetic energy integrals are just fixed linear combinations of overlap integrals of higher angular momentum, the same argument holds for them as well.

For $S$ orbitals, the two-electron Coulomb integral may be written as

$$h_{acbd} = \frac{S_{ab} S_{cd}}{R_{PQ}} \operatorname{erf}(\sqrt{\alpha} R_{PQ}) \tag{9.15}$$

where erf is the error function, $P$ and $Q$ are Gaussian centers formed through application of the Gaussian product rule to $|G_a\rangle |G_b\rangle$ and $|G_c\rangle |G_d\rangle$ respectively. $R_{PQ}$ is the distance between the two Gaussian centers $P$ and $Q$ and $\alpha$ is the reduced exponent derived from $P$ and $Q$. For clarity, this may be bounded by the simpler expression

$$h_{acbd} \leq \min\left( \frac{4\alpha}{\pi} S_{ab} S_{cd}, \frac{S_{ab} S_{cd}}{R_{PQ}} \right) \tag{9.16}$$

The first of these two expressions in the min function comes from the short range bound and the latter from the long range bound of the error function. These bounds show that the integrals are determined by products of overlap terms, such that in the regime where overlap integrals scale linearly, we expect $O(M^2)$ significant two-electron terms. Moreover, as seen in the long range bound of the two-electron integral, there is some further asymptotic distance beyond which these interactions may be completely neglected, yielding an effectively linear scaling number of significant integrals. This limit can be quite large however, thus practically one expects to observe a quadratic scaling in the number of two-electron integrals (TEI).

Additionally, we note from the form of the integrals, that the maximal values the two-electron integrals will attain are determined by the basis set parameters, such as the width of the Gaussian basis functions or their angular momentum. The implication of this, is that the maximal integral magnitude for the four index two-electron integrals, $|h^{\mathrm{TEI}}_{\mathrm{max}}|$ will be independent of the molecular size for standard atom centered Gaussian basis sets, and may be treated as a constant for scaling analysis that examine cost as a function of physical system size with fixed chemical composition. The overlap and kinetic energy integrals will similarly have a maximum independent of molecular size past a very small length scale. However, the nuclear attraction integrals must also be considered.

While not typically considered a primary source of difficulty due to the relative ease of evaluation with respect to two-electron integrals, we separate the nuclear attraction integrals here due to the fact that the maximal norm of the elements may change as well. The nuclear attraction matrix element between $S$ functions may be written as

$$h^{\mathrm{nuc}}_{ab} = -\sum_i \frac{Z_i S_{ab}}{R_{Pi}} \operatorname{erf}(\sqrt{p} R_{Pi}) \tag{9.17}$$

where $Z_i$ is the nuclear charge and $R_{Pi}$ refers to the distance between the Gaussian center

$P$ with total exponent $p$ formed from the product $|G_a\rangle |G_b\rangle$ to the position of the $i$'th nuclei. Following from the logic above, from the exponentially vanishing overlap $S_{ab}$, at some distance, we expect only a linear number of these integrals to be significant. However, each of the integrals considers the sum over all nuclei, which can be related linearly to the number of basis functions in atom centered Gaussian basis sets. Thus the maximal one-electron integral is not a constant, but rather can be expected to scale with the Coulomb sum over distant nuclear charges. A conservative bound can be placed on such a maximal element as follows.

The temperature and pressure a molecule reside in will typically determine the minimal allowed separation of two distinct nuclei, and will thus define a maximum nuclear density $\rho_{\max}$. Denote the maximum nuclear charge in the systems under consideration as $Z_{\max}$. The maximal density and the number of nuclei will also define a minimal radius that a sphere of charge may occupy $r_{\max}$,

$$r_{\max}^3 = \frac{3 Z_{\max} N_{\text{nuc}}}{4\pi\rho_{\max}} \tag{9.18}$$

where $N_{\text{nuc}}$ is the number of nuclei in the system. Modeling the charge as spread uniformly within this minimal volume and using the maximum of the error function to find a bound on the maximum for the nuclear attraction matrix element, we find

$$\begin{aligned} |h_{ab}^{\text{nuc}}| &< 4\pi\rho_{\max} S_{ab} \left| \int_0^{r_{\max}} r^2 dr \frac{1}{r} \right| \\ &= 2\pi\rho_{\max} S_{ab} r_{\max}^2 \\ &= \beta_{ab} N_{\text{nuc}}^{2/3} \end{aligned} \tag{9.19}$$

where $\beta_{ab}$ is now a system size independent quantity determined only by basis set parameters at nuclei $a$ and $b$, and the size dependence is bounded as $O(N_{\text{nuc}}^{2/3})$. Atom centered Gaussian basis sets will have a number of a basis functions which is a linear multiple of the number

of nuclei, and as such we may now bound the maximal one-electron integral (OEI) element as

$$|h_{\max}^{\mathrm{OEI}}| < \beta_{\max}^{\mathrm{OEI}} M^{2/3} \tag{9.20}$$

### 9.1.4 Effect of truncation

The above analysis demonstrates that given some integral magnitude threshold, $\delta$, there exists a characteristic distance $d$ between atomic centers, beyond which integrals may be neglected. If one is interested in a total precision $\epsilon$ in the energy $E_i$, it is important to know how choosing $\delta$ will impact the solution, and what choice of $\delta$ allows one to retain a precision $\epsilon$.

By specification, the discarded integrals are small with respect to the rest of the Hamiltonian (sometimes as much as 10 orders of magnitude smaller in standard calculations). As such, one expects a perturbation analysis to be accurate. Consider the new, truncated Hamiltonian $H_t = H + V$, where $V$ is the negation of the sum of all removed terms, each of which have magnitude less than $\delta$.

Assuming a non-degenerate spectrum for $H$, from perturbation theory we expect the leading order change in eigenvalue $E_i$ to be given by

$$\Delta E_i = \langle \Psi_i | \, V \, | \Psi_i \rangle \tag{9.21}$$

if the number of terms removed from the sum is given by $N_r$, a worst case bound on the magnitude of this deviation follows from the spectrum of the creation and annihilation operators and is given by

$$|\Delta E_i| \leq \sum_{\{h_i : |h_i| < \delta\}} |h_i| \leq N_r \delta \tag{9.22}$$

where $\{h_i : |h_i| < \delta\}$ is simply the set of Hamiltonian elements with norm less than $\delta$ and the first inequality follows directly from the triangle inequality. We emphasize that this is

a worst case bound, and generically one expects at least some cancellation between terms, such as kinetic and potential terms, when the Hamiltonian is considered as a whole. Some numerical studies of these cancellation effects have been performed [224], but additional studies are required. Regardless, under this maximal error assumption, by choosing a value

$$\delta \leq \frac{\epsilon}{N_r} \tag{9.23}$$

one retains an accuracy $\epsilon$ in the final answer with respect to the exact answer when measuring the eigenvalue of the truncated Hamiltonian $H_t$. Alternative, one may use the tighter bound based on the triangle inequality and remove the maximum number of elements such that the total magnitude of removed terms is less than $\epsilon$. From the looser but simpler bound, we see a reduction of scaling from $M^4$ to $M^2$ would require removal of the order of $M^4$ terms from the Hamiltonian, this constraint on $\delta$ can be rewritten in terms of $M$ as

$$\delta \leq \frac{\epsilon}{M^4} \tag{9.24}$$

While the perturbation of the eigenvalue will have a direct influence on energy projective measurement methods such as quantum phase estimation, other methods evaluate the energy by averaging. In this case, we do not need to appeal to perturbation theory, and the $\delta$ required to achieve a desired $\epsilon$ can be found directly.

$$\langle H_t \rangle = \langle \Psi_i | \, H_t \, | \Psi_i \rangle \tag{9.25}$$

$$= E_i + \langle \Psi_i | \, V \, | \Psi_i \rangle \tag{9.26}$$

We find that under our assumption of worst case error for averaging, the result is identical to that of the first order perturbation of the eigenvalue $E_i$,

$$|\Delta \langle H_t \rangle| \leq \sum_{\{h_i : |h_i| < \delta\}} |h_i| \leq N_r \delta \tag{9.27}$$

In summary, we find that for both the consideration of the ground state eigenvalue and the average energy of the ground state eigenvector, there is a simple formula for the value of $\delta$, which scales polynomially in the system size, below which one may safely truncate to be guaranteed an accuracy $\epsilon$ in the final answer. Moreover it suggests a simple strategy that one may utilize to achieve the desired accuracy. That is, sort the integrals in order of magnitude, and remove the maximum number of integrals such that the total magnitude of removed integrals is less than $\epsilon$.

On the subject of general truncation, we note that while there exist may Hamiltonians with the same structure as the second quantized electronic structure Hamiltonian that have the property that removal of small elements will cause a drastic shift in the character of the ground state, this has not been seen for physical systems in quantum chemistry. In practice it is observed that removing elements on the order of $\delta = 10^{-10}$ and smaller is more than sufficient to retain both qualitative and quantitative accuracy in systems of many atoms [10, 290, 145, 11]. Moreover, the convergence with respect to this value may be tested easily for any systems under consideration.

### 9.1.5 Onset of favorable scaling

While the above analysis shows that locality of interactions in local basis sets provides a promise that beyond a certain length scale, the number of non-negligible integrals will scale quadratically in the number of basis functions, it does not provide good intuition for the size of that length scale in physical systems of interest. Here we provide numerical examples for chemical systems in basis sets used so far in quantum computation for quantum chemistry. The precise distance at which locality starts to reduce the number of significant integrals depends, of course, on the physical system and the basis set used. In particular, larger, more

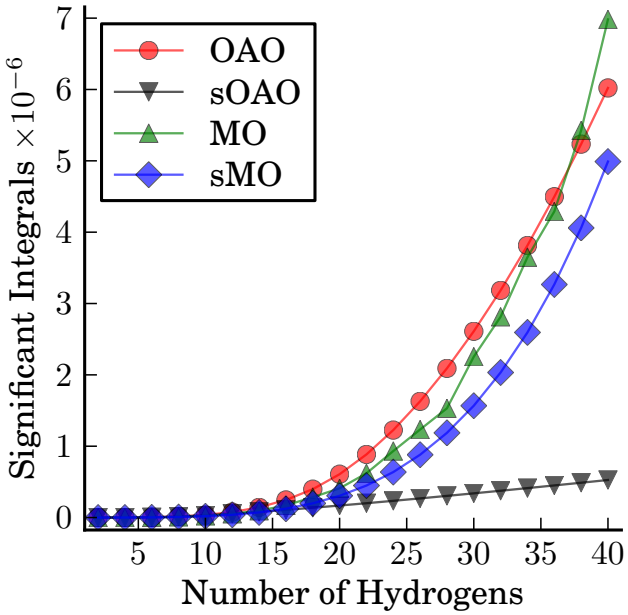



Figure 9.1: The number of significant(magnitude $> 10^{-15}$) spin-orbital integrals in the STO-3G basis set as a function of the number of hydrogens in a linear hydrogen chain with a separation of 1 $a_0$ for the Hartree-Fock canonical molecular orbital basis(MO) and the symmetrically orthogonalized atomic orbital basis (OAO). The sMO and sOAO, shows the same quantity with a sharper cutoff ($10^{-7}$) and demonstrates the advantage to localized atomic basis functions at length scales as small as 10 Å.

diffuse basis sets are known to exhibit these effects at comparatively larger length scales than minimal, compact basis sets. However the general scaling arguments given above hold for all systems of sufficient size.

An additional consideration which must be made for quantum computation, is that as of yet, no general technology has been developed for direct simulation in non-orthogonal basis sets. This prohibits direct simulation in the bare atomic orbital basis, however the use of Löwdin symmetric orthogonalization yields the orthogonal basis set closest to the original atomic orbital basis set in an $l^2$ sense [185, 194]. We find that this is sufficient for the systems we consider, but note that there have been a number of advances in orthogonal basis sets that are local in both the occupied and virtual spaces and may find utility in quantum

computation [300]. Moreover, there has been recent work in the use of multiresolution wavelet basis sets that have natural sparsity and orthogonality while providing provable error bounds on the choice of basis [137]. Such a basis also allows one to avoid costly integral transformations related to orthogonality, which are known to scale as $O(M^5)$ when performed exactly. Further research is needed to explore the implications for quantum computation with these basis sets, and we focus here on the more common atom-centered Gaussian basis sets.

As a prototype system, we consider chains of hydrogen atoms separated by 1 Bohr ($a_0$) in the STO-3G basis set, an artificial system that can exhibit a transition to a strongly correlated wavefunction [127]. We count the total number of significant integrals for values of $\delta$ given by $10^{-15}$ and $10^{-7}$ for the symmetrically orthogonalized atomic orbital (OAO) basis and the canonical Hartree-Fock molecular orbital (MO) basis. The results are displayed in Fig. 9.1 and demonstrate that with a cutoff of $\delta = 10^{-7}$ the localized character of the OAO's allows for a savings of on the order of $6 \times 10^6$ integrals with respect to the more delocalized canonical molecular orbitals. The s in the labeling of the orbital bases simply differentiates between two possible cutoffs. These dramatic differences begin to present with atomic chains as small as 10 Å in length in this system with this basis set.

As an additional example, we consider linear alkane chains of increasing length. The results are displayed in Fig. 9.2 and again display the dramatic advantages of preserving locality in the basis set. By the point one reaches 10 carbon atoms, a savings of almost $10^8$ integrals can be achieved at a truncation level of $10^{-7}$.

Although localized basis sets provide a definitive scaling advantage in the medium-large size limit for molecules, one often finds that in the small molecule limit canonical molecular orbitals, the orbitals from the solution of the Hartree-Fock equations under the canonical

condition, provide a more sparse representation. This is observed in Figs 9.1 and 9.2 for the smallest molecule sizes, and the transition for this behavior will generally be basis set dependent. The reason is that at smaller length scales, the "delocalized" canonical molecule orbitals have similar size to the more localized atomic orbitals, but with the additional constraint of the canonical condition, a sufficient but not necessary condition for the solution of the Hartree-Fock equations that demands the Fock matrix be diagonal (as opposed to the looser variational condition of block-diagonal between the occupied and virtual spaces). A side effect of the canonical condition is that in the canonical molecular orbital basis many of the $h_{pqrs}$ terms for distinct indices are reduced in magnitude. However, there are not enough degrees of freedom present in the orbital rotations for this effect to persist to larger length scales, and as a result local basis sets eventually become more advantageous. Moreover, it is known that at larger length scales, the canonical conditions tend to favor maximally delocalized orbitals, which can reduce the advantages of locality. These effects have been studied in some detail in the context of better orbital localizations by relaxing the canonical condition in Hartree-Fock and the so-called Least-Change Hartree-Fock method coupled with fourth-moment minimization [300].

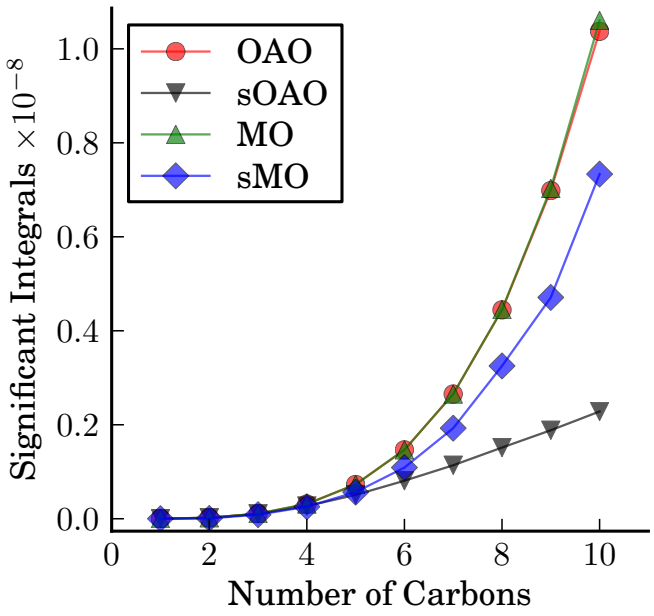



Figure 9.2: The number of significant(magnitude $> 10^{-15}$) spin-orbital integrals in the STO-3G basis set as a function of the number of carbons in a linear alkane chain for the Hartree-Fock canonical molecular orbital basis(MO) and the symmetrically orthogonalized atomic orbital basis (OAO). The sMO and sOAO shows the same quantity with a sharper cutoff ($10^{-7}$) and demonstrates the dramatic advantage to localized atomic basis even at this small atomic size.

## 9.2 Quantum energy estimation

Almost all algorithms designed for the study of quantum chemistry eigenstates on a quantum computer can be separated into two distinct parts: 1. state preparation and 2. energy estimation. For the purposes of analysis, it is helpful to treat the two issues separately, and in this section we make the standard assumption in doing so, that an oracle capable of producing good approximations to the desired eigenstates $|\Psi_i\rangle$ at unit cost is available. Under this assumption, energy estimation for a fixed desired precision $\epsilon$ is known to scale polynomially in the size of the system for quantum chemistry, however the exact scaling costs and tradeoffs depend on the details of the method used. Here we compare the costs and benefits of two prominent methods of energy estimation used in quantum computation for chemistry: quantum phase estimation and Hamiltonian averaging.

### 9.2.1 Quantum phase estimation

The first method used for the energy estimation of quantum chemical states on a quantum computer was quantum phase estimation [169, 2, 11]. The method works by evolving the give quantum eigenstate $|\Psi_i\rangle$ forward under the system Hamiltonian $H$ for a time $T$, and reading out the accumulated phase, which can be easily mapped to the associated eigenenergy $E_i$. While the basic algorithm and its variations can have many different components, the cost is universally dominated by the coherent evolution of the system.

To evolve the system under the Hamiltonian, one must find a scalable way to implement the unitary operator $U = e^{-iHT}$. The standard procedure for accomplishing this task is the use of Suzuki-Trotter splitting [266, 259], which approximates the unitary operator(at first

order) as

$$U = e^{-iHT} = \left(e^{-iH(T/m)}\right)^m = \left(e^{-i(\sum_i H_i)\Delta t}\right)^m \approx \left(\prod_i e^{-iH_i\Delta t}\right)^m \tag{9.28}$$

where $\Delta t = T/m$ and $H_i$ is a single term from the Bravyi-Kitaev transformed system Hamiltonian. Higher order Suzuki-Trotter operator splittings and their benefits have been studied in the context of quantum simulation [33], but we largely focus on the first order formula in this work. If each of the simpler unitary operators $e^{-iH_i\Delta t}$ has a known gate decomposition, the total time evolution can be performed by chaining these sequences together.

The use of the Suzuki-Trotter splitting can be thought of as an evolution under an approximate Hamiltonian $\tilde{H}$, given by $e^{-i\tilde{H}T}$, whose eigenspectrum deviates from the original Hamiltonian by a factor depending on time-step $\Delta t$. The precise dependence of this bias depends on the order of the Suzuki-Trotter expansion used. The total resolution, $\epsilon$, in the energies of the approximate Hamiltonian $\tilde{H}$ is determined by the total evolution time $T$. Thus to achieve an accuracy of $\epsilon$ in the final energy, one must utilize a time step $\Delta t$ small enough that the total bias is less than $\epsilon$ and a total run time $T$ such that the resolution is better than $\epsilon$. If the number of gates required to implement a single timestep $\Delta t$ is given by $N_g$, then the dominant cost of simulation (all of which must be done coherently) is given by

$$N_c = N_g\lceil\frac{T}{\Delta t}\rceil \tag{9.29}$$

The total evolution time $T$ required to extract an eigenvalue to chemical precision $\epsilon_{\text{chem}} = 10^{-3}$ is typically set at the Fourier limit independent of molecular size and thus can be considered a constant for scaling analysis. We then focus on the number of gates per Suzuki-Trotter time step, $N_g$, and the time step $\Delta t$ required to achieve the desired precision.

In a first order Suzuki-Trotter splitting, the number of gates per Trotter time step is given by the number of terms in the Hamiltonian multiplied by the number of gates required to implement a single elementary term for the form $e^{-iH_i\Delta t}$. The gates per elementary term can vary based on the particular integral, however for simplicity in developing bounds we consider this as constant here. The number of terms, is known from previous analysis in this work to scale as $O(M^2)$ or in the truly macroscopic limit $O(M)$. The number of gates required to implement a single elementary term depends on the transformation used from fermionic to qubit operators. The Jordan-Wigner transformation [156] results in non-local terms that carry with them an overhead that scales as the number of qubits, which in this case will be $O(M)$. Although there have been developments in methods to use teleportation to perform these non-local operations in parallel [77] and by improving the efficiency of the circuits computing the phases in the Jordan-Wigner transformation [140], these issues can also be alleviated by choosing the Brayvi-Kitaev transformation that carries an overhead only logarithmic in the number of qubits, $O(\log M)$ [56, 243]. As a result, one expects the number of gates per Suzuki-Trotter time step $N_g$ to scale as $O(M^2 \log M)$ or in a truly macroscopic limit $O(M \log M)$.

To complete the cost estimate with fixed total time $T$, one must determine how the required time step $\Delta t$ scales with the size of the system. As mentioned above, the use of the Suzuki-Trotter decomposition for the time evolution of $H$ is equivalent to evolution under an effective Hamiltonian $\tilde{H} = H + V$, where the size of the perturbation is determined by the order of the Suzuki-Trotter formula used and the size of the timestep. Once the order of the Suzuki-Trotter expansion to be used has been determined, the requirement on the timestep is such that the effect of $V$ on the eigenvalue of interest is less than the desired accuracy in the final answer $\epsilon$.

This has been explored previously [140, 224], but we now examine this scaling in our context. To find $V$, one may expand the $k$'th order Suzuki-Trotter expansion of the evolution of $\tilde{H}$ into a power series as well as the power series of the evolution operator $\exp\left[-i\left(H+V\right)\Delta t\right]$, and find the leading order term $V$. As a first result, we demonstrate that for a $k$'th order propagator, the leading perturbation on the ground state eigenvalue for a non-degenerate system is $O(\Delta t)^{k+1}$.

Recall the power series expansion for the propagator

$$\exp\left[-i\left(H+V\right)\Delta t\right]=\sum_{j=0}^{\infty}\frac{(-i)^{j}}{j!}\left(H+V\right)^{j}\left(\Delta t\right)^{j} \tag{9.30}$$

The definition of a $k$'th order propagator, is one is that correct through order $k$ in the power series expansion. As such, when this power series is expanded, $V$ must make no contribution in the terms until $O((\Delta t)^{k+1})$. For this to be possible, it's clear that $V$ must depend on $\Delta t$. In order for it to vanish for the first $k$ terms, $V$ must be proportional to $(\Delta t)^{k}$. Moreover, due to the alternation of terms between imaginary and real at each order in the power series with the first term being imaginary, the first possible contribution is order $(\Delta t)^{k}$ and imaginary. As is common in quantum chemistry, we assume a non-degenerate and real ground state, and thus the contribution to the ground state eigenvalue is well approximated by first order perturbation theory as

$$E^{(1)}=\left\langle \Psi_{g}\right|V\left|\Psi_{g}\right\rangle \tag{9.31}$$

however, as $V$ is imaginary Hermitian and the ground state is known to be real in quantum chemistry, this expectation value must vanish. Thus the leading order perturbation to the ground state eigenvalue is at worst the real term depending on $(\Delta t)^{k+1}$.

To get a more precise representation of $V$ for a concrete example, we now consider the first order ($k=1$) Suzuki-Trotter expansion. As expected, the leading order imaginary

error term is found to be

$$V^{(0)} = \frac{\Delta t}{2} \sum_{j<k} i \left[ H_j, H_k \right] \tag{9.32}$$

whose contribution must vanish due to it being an imaginary Hermitian term. Thus we look to the leading contributing error depending on $(\Delta t)^2$, which has been obtained previously[224] from a Baker-Campbell-Hausdorff(BCH) expansion to read

$$V^{(1)} = \frac{(\Delta t)^2}{12} \sum_{i \leq j} \sum_{j} \sum_{k<j} \left[ H_i \left( 1 - \frac{\delta_{ij}}{2} \right), \left[ H_j, H_k \right] \right] \tag{9.33}$$

Thus the leading order perturbation is given by third powers of the Hamiltonian operators. To proceed, we count the number of one- and two-electron integrals separately as $N_{\text{int}}^{\text{OEI}}$ and $N_{\text{int}}^{\text{TEI}}$ respectively. Their maximal norm elements are similarly denoted by $h_{\text{max}}^{\text{OEI}}$ and $h_{\text{max}}^{\text{TEI}}$. From this, we can draw a worst case error bound on the perturbation of the eigenvalue given by

$$\begin{aligned} E^{(1)} &\leq \frac{(\Delta t)^2}{12} \sum_{i \leq j} \sum_{j} \sum_{k<j} \left| H_i \left( 1 - \frac{\delta_{ij}}{2} \right), \left[ H_j, H_k \right] \right| \\ &\leq \left( |h_{\text{max}}^{\text{OEI}}| N_{\text{int}}^{OEI} + |h_{\text{max}}^{\text{TEI}}| N_{\text{int}}^{\text{TEI}} \right)^3 (\Delta t)^2 \\ &\leq \left( |\beta_{\text{max}}^{\text{OEI}}| M^{2/3} N_{\text{int}}^{\text{OEI}} + |h_{\text{max}}^{\text{TEI}}| N_{\text{int}}^{\text{TEI}} \right)^3 (\Delta t)^2 \end{aligned} \tag{9.34}$$

Where the first inequality follows from the triangle inequality and the second is a looser, but simpler bound, that may be used to elucidate the scaling behavior. Holding the looser bound to the desired precision in the final answer $\epsilon$, this yields

$$\Delta t \leq \left[ \frac{\epsilon}{\left( |\beta_{\text{max}}^{\text{OEI}}| M^{2/3} N_{\text{int}}^{\text{OEI}} + |h_{\text{max}}^{\text{TEI}}| N_{\text{int}}^{\text{TEI}} \right)^3} \right]^{1/2} \tag{9.35}$$

We emphasize that this is a worst case bound, including no possible cancellation between Hamiltonian terms. Some preliminary work has been done numerically in establishing average cancellation between terms that shows these worst case bounds are too

pessimistic [224]. Continuing, we expect the total scaling under a first order Suzuki-Trotter expansion using a Bravyi-Kitaev encoding to be bounded by

$$N_c = N_g \lceil \frac{T}{\Delta t} \rceil \leq \frac{N_g}{\epsilon \Delta t} \leq \frac{\left( |\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}} \right)^{3/2} N_{\mathrm{int}} \log M}{\epsilon^{3/2}} \tag{9.36}$$

and in the large size limit where the number of significant two-electron integrals in a local basis set scales quadratically and the number of significant one-electron integrals scales linearly, this may be bounded by

$$N_c \leq \kappa \frac{\left( |\beta_{\max}^{\mathrm{OEI}}| M^{5/3} + |h_{\max}^{\mathrm{TEI}}| M^2 \right)^{3/2} (M^2 + M) \log M}{\epsilon^{3/2}} \tag{9.37}$$

where $\kappa$ is a positive constant that will depend on the basis set and this expression scales as $O(M^5 \log M)$ in the number of spin-orbital basis functions.

### 9.2.2 Hamiltonian averaging

The quantum phase estimation algorithm has been central in almost all algorithms for energy estimation in quantum simulation. However, it has a significant practical drawback in that after state preparation, all the desired operations must be performed coherently. A different algorithm for energy estimation has recently been introduced [223, 296] that lifts all but an $O(1)$ coherence time requirement after state preparation, making it amenable to implementation on quantum devices in the near future. We briefly review this approach, which we will call Hamiltonian averaging, and bound its costs in applications for quantum chemistry.

As in quantum phase estimation, in Hamiltonian averaging one assumes the eigenstates $|\Psi_i\rangle$ are provided by some oracle. By use of either the Jordan-Wigner or Bravyi-Kitaev transformation, the Hamiltonian may be written as a sum of tensor products of Pauli

operators. These transformations at worst conserve the number of independent terms in the Hamiltonian, thus we may assume for our worst case analysis the number of terms is fixed by $N_{int}$ and the coefficients remain unchanged. From the provided copy of the state and transformed Hamiltonian, to obtain the energy one simply performs the average

$$\langle H \rangle = \sum_{i,j,k,... \in x,y,z} h_{ijk...} \langle \sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k ... \rangle \tag{9.38}$$

by independent Pauli measurements on the provided state $|\Psi_i\rangle$ weighted by the coefficients $h_{ijkl...}$, which are simply a relabeling of the previous two-electron integrals for convenience with the transformed operators. As $|\Psi_i\rangle$ is an eigenstate, this average will correspond to the desired eigenvalue $E_i$ with some error related to sampling that we now quantify.

Consider an individual term

$$X_{ijkl...} = h_{ijkl...} \sigma_1^i \otimes \sigma_2^j \otimes \sigma_3^k ... \tag{9.39}$$

it is clear from the properties of qubit measurements, that the full range of values this quantity can take on is $[-h_{ijkl...}, h_{ijkl...}]$. As a result, we expect that the variance associated with this term can be bounded by

$$\text{Var}\left[X_{ijkl...}\right] \leq |h_{ijkl...}|^2 \tag{9.40}$$

Considering a representative element, namely the maximum magnitude integral element $h_{\max}$, we can bound the variance of $H$ as

$$\text{Var}\left[H\right] \leq N_{int}^2 |h_{\max}|^2 \tag{9.41}$$

The variance of the mean, which is the relevant term for our sampling error, comes from the central limit theorem and is bounded by

$$\text{Var}\left[\langle H \rangle\right] \leq \frac{\text{Var}\left[H\right]}{N} \tag{9.42}$$

where $N$ is the number of independent samples taken of $\langle H \rangle$. Collecting these results, we find

$$\mathrm{Var}\left[\langle H \rangle\right] \leq \sum \frac{|h_{ijkl\ldots}|^2}{N} \leq \frac{\left(|\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^2}{N} \tag{9.43}$$

Now setting the variance to the desired statistical accuracy $\epsilon^2$ (which corresponds to a standard error of $\epsilon$ at a 68% confidence interval), we find the number of independent samples expected, $N_s$, is bounded by

$$N_s \leq \frac{\left(|\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^2}{\epsilon^2} \tag{9.44}$$

If a single independent sample of $\langle H \rangle$ requires the measurement of each of the $N_{int}$ quantities, then the bound on the total cost in the number of state preparations and measurements, $N_m$ is

$$N_m \leq \frac{N_{int}\left(|\beta_{\max}^{\mathrm{OEI}}| M^{2/3} N_{\mathrm{int}}^{\mathrm{OEI}} + |h_{\max}^{\mathrm{TEI}}| N_{\mathrm{int}}^{\mathrm{TEI}}\right)^2}{\epsilon^2} \tag{9.45}$$

which if one considers the large size limit,such that the number of two-electron integrals scales quadratically and the number of one-electron integrals scales linearly, we find

$$N_m \leq \kappa \frac{(M + M^2)\left(|\beta_{\max}^{\mathrm{OEI}}| M^{5/3} + |h_{\max}^{\mathrm{TEI}}| M^2\right)^2}{\epsilon^2} \tag{9.46}$$

where $\kappa$ is a positive constant that depends upon the basis set. It is clear that this expression scales as $O(M^6)$ in the number of spin-orbital basis functions. We see from this, that under the same maximum error assumptions, Hamiltonian averaging scales only marginally worse in the number of integrals and precision as compared to quantum phase estimation performed with a first order Suzuki-Trotter expansion, but has a coherence time requirement of $O(1)$ after each state preparation. Note that each measurement is expected to require single qubit rotations that scale as either $O(M)$ for the Jordan-Wigner transformation or

$O(\log M)$ for the Bravyi-Kitaev transformation. However, we assume that these trivial single qubit rotations can be performed in parallel independent of the size of the system without great difficulty, and we thus don't consider this in our cost estimate. This method is a suitable replacement for quantum phase estimation in situations where coherence time resources are limited and good approximations to the eigenstates are readily available. Additional studies are needed to quantify the precise performance of the two methods beyond worst case bounds.

## 9.3 Using imperfect oracles

A central assumption for successful quantum phase estimation and typically any energy evaluation scheme is access to some oracle capable of producing good approximations to the eigenstate of interest, where a "good" approximation is typically meant to imply an overlap that is polynomial in the size of the system. Additionally, a supposed benefit of phase estimation over Hamiltonian averaging is that given such a good (but not perfect) guess, by projective measurement in the energy basis, in principle one may avoid any bias in the final energy related to the initial state. Here we examine this assumption in light of the Van-Vleck catastrophe [270], which we review below, and examine the consequences for measurements of the energy by QPE and Hamiltonian averaging.

The Van Vleck catastrophe [270] refers to an expected exponential decline in the quality of trial wavefunctions, as measured by overlap with the true wavefunction of a system, as a function of size. We study a simple case of the catastrophe here in order to frame the consequences for quantum computation. Consider a model quantum system consisting of a collection of $N$ non-interacting two level subsystems with subsystem Hamiltonians given by $H_i$. These subsystems have ground and excited eigenstates $|\psi_g^i\rangle$ and $|\psi_e^i\rangle$ with eigenenergies

$E_g < E_e$, such that the total Hamiltonian is given by

$$H = \sum_i H_i \tag{9.47}$$

and eigenstates of the total Hamiltonian are formed from tensor products of the eigenstates of the subsystems. As such the ground state of the full system is given by

$$|\Psi_g\rangle = \bigotimes_{i=0}^{N-1} |\psi_g^i\rangle \tag{9.48}$$

Now suppose we want to measure the ground state energy of the total system, but the oracle is only capable of producing trial states for each subsystem $|\psi_t^i\rangle$ such that $\langle\psi_t^i|\psi_g^i\rangle = \Delta$, where $|\Delta| < 1$. The resulting trial state for the whole system is

$$|\Psi_t\rangle = \bigotimes_{i=0}^{N-1} |\psi_t^i\rangle \tag{9.49}$$

From normalization of the two level system, we may also write the trial state as

$$|\psi_t^i\rangle = \Delta\,|\psi_g^i\rangle + e^{-i\theta}\sqrt{1-\Delta^2}\,|\psi_e^i\rangle \tag{9.50}$$

where $\theta \in [0, 2\pi)$. Moreover, from knowledge of the gap, one can find the expected energy on each subsystem, which is given by

$$\langle\psi_t^i|\,H_i\,|\psi_t^i\rangle = \Delta^2 E_g + (1-\Delta^2)E_e \tag{9.51}$$

For the case of Hamiltonian averaging on the total system, the expected answer is given by

$$\begin{aligned} E &= \langle\Psi_t|\,H\,|\Psi_t\rangle \\ &= \sum_{i=0}^{N-1} \langle\psi_t^i|\,H_i\,|\psi_t^i\rangle \\ &= N(\Delta^2 E_g + (1-\Delta^2)E_e) \end{aligned} \tag{9.52}$$

which yields an energy bias from the true ground state, $\epsilon_b$, given by

$$\begin{aligned}\epsilon_b &= N(\Delta^2 E_g + (1-\Delta^2)E_e) - NE_g \\ &= N(1-\Delta^2)(E_e - E_g) \\ &= N(1-\Delta^2)\omega \end{aligned} \tag{9.53}$$

where we denote the gap for each subsystem as $\omega = (E_e - E_g)$. As such, it is clear that the resulting bias is only linear in the size of the total system $N$.

Quantum phase estimation promises to remove this bias by projecting into the exact ground state. However, this occurs with a probability proportional to the square of the overlap of the input trial state with the target state. In this example, this is given by

$$|\langle \Psi_t | \Psi_g \rangle| = |\Delta|^{2N} \tag{9.54}$$

which is exponentially small in the size of the system. That is, quantum phase estimation is capable of removing the bias exactly in this example non-interacting system, but at a cost which is exponential in the size of the system. The expected cost of removing some portion of the bias may be calculated by considering the distribution of states and corresponding energies.

Consider first the probability of measuring an energy with a bias of $\epsilon(M) = M(1-\Delta^2)\omega$. For this to happen, it is clear that exactly $M$ of the subsystems in the measured state are in the excited state. It is clear that this is true for $\begin{pmatrix} N \\ M \end{pmatrix}$ eigenstates, and the square of the overlap with such an eigenstate is $\left(\Delta^2\right)^{N-M}\left(1-\Delta^2\right)^M$ or

$$P(\epsilon(M)) = \begin{pmatrix} N \\ M \end{pmatrix} \left(\Delta^2\right)^{N-M}\left(1-\Delta^2\right)^M \tag{9.55}$$

which is clearly a binomial distribution. As a result, in the large $N$ limit, this distribution is well approximated by a Gaussian and we may write

$$P(\epsilon(M)) \approx \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2}\left(\frac{M-\bar{N}}{\sigma}\right)^2\right] \tag{9.56}$$

$$\bar{N} = N(1-\Delta^2) \tag{9.57}$$

$$\sigma^2 = N\Delta^2(1-\Delta^2) \tag{9.58}$$

Bringing this together, we find that the probability of measuring a bias of less than $\epsilon(M)$ is given by

$$\begin{aligned} P(<\epsilon(M)) &= \frac{1}{\sqrt{2\pi\sigma^2}} \int_0^M dM' \exp\left[-\frac{1}{2}\left(\frac{M'-\bar{N}}{\sigma}\right)^2\right] \\ &= \frac{1}{2}\left[\mathrm{erf}\left(\frac{M-\bar{N}}{\sqrt{2\sigma^2}}\right) + \mathrm{erf}\left(\frac{\bar{N}}{\sqrt{2\sigma^2}}\right)\right] \end{aligned} \tag{9.59}$$

where erf is again the error function.

Thus the expected cost in terms of number of repetitions of the full phase estimation procedure to remove a bias of at least $\epsilon(M)$ from this model system is

$$\begin{aligned} C(<\epsilon(M)) &= \frac{1}{P(<\epsilon(M))} \\ &= 2\left[\mathrm{erf}\left(\frac{M-\bar{N}}{\sqrt{2\sigma^2}}\right) + \mathrm{erf}\left(\frac{\bar{N}}{\sqrt{2\sigma^2}}\right)\right]^{-1} \end{aligned} \tag{9.60}$$

We plot the expected cost function for a range of oracle guess qualities $\Delta$ on a modest system of $N = 100$ in Fig 9.3. From this, we see that the amount of bias that can feasibly be removed depends strongly on the quality of the oracle guess. Generically, we see that for any fixed imperfect guess on the subsystem level($|\Delta| < 1$), there will be an exponential cost in phase estimation related to perfect removal of the bias.

This problem can be circumvented by improving the quality of the subsystem guesses as a function of system size. In particular, one can see that if $|\Delta|$ is improved as $(1-1/(2N))$

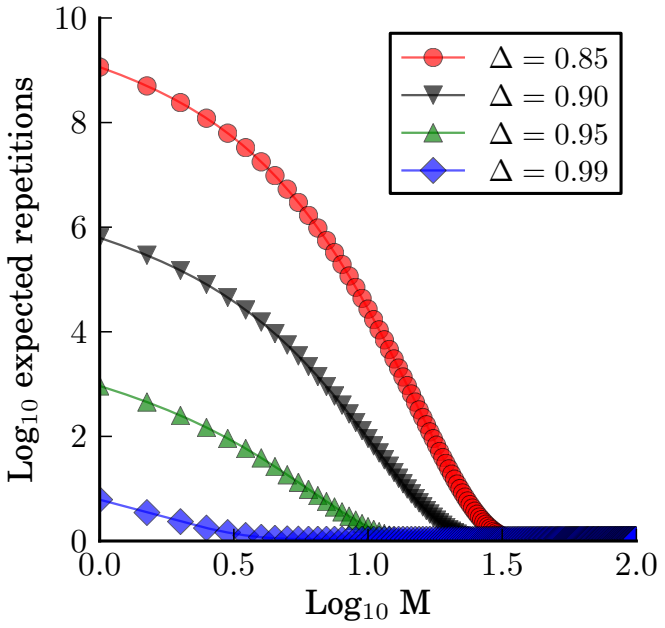



Figure 9.3: A log log plot of the expected cost in number of repetitions of measuring an energy with a bias $\epsilon(M)$ as a function of $M$ in quantum phase estimation for different values of the oracle quality $\Delta$. A system of $N = 100$ non-interacting subsystems is considered. A perfect, unbiased answer corresponds to $M = 0$ with expected cost $O(\Delta^{2N})$, however to aid in visualization this plot is provided only beyond $M = 1$. In general one sees that depending on the oracle quality $\Delta$, different fractions of the bias may be removed with ease, but there is always some threshold for imperfect guesses ($|\Delta| < 1$) such that there is an exponential growth in cost.

then $|\Delta|^{2N}$ is $O(1)$. However, as the subsystems in a general case could be of arbitrary size, classical determination of a subsystem state of sufficient quality may scale exponentially in the required precision and thus system size. Moreover, one would not expect the problem to be easier in general cases where interactions between subsystems are allowed. As a result, further developments in variational methods [223], quantum cooling [292], and adiabatic state preparation [11, 14, 272] will be of key importance in this area. Moreover improvements in the ansatze used to prepare the wave function such as multi-configurational self consistent field calculations(MCSCF) [277, 272] or unitary coupled cluster (UCC) [296] will be integral parts of any practical quantum computing for quantum chemistry effort.

## 9.4 Adiabatic computation

A complementary solution for the problem of molecular simulation on quantum computers is that of adiabatic quantum computation. It is not known to show the same direct dependence on the overlap of the initial guess state as QPE, which may allow it to solve different problems than the quantum phase estimation or variational quantum eigensolver in practice. In [14], Babbush *et al.* show how to scalably embed the eigenspectra of molecular Hamiltonians in a programmable physical system so that the adiabatic algorithm can be applied directly. In this scheme, the molecular Hamiltonian is first written in second quantization using fermionic operators. This Hamiltonian is then mapped to a qubit Hamiltonian using the Bravyi-Kitaev transformation [56, 243]. The authors show that the more typical Jordan-Wigner transformation cannot be used to scalably reduce molecular Hamiltonians to 2-local qubit interactions as the Jordan-Wigner transformation introduces linear locality overhead which translates to an exponential requirement in the precision of the couplings when perturbative gadgets are applied. Perturbative gadgets are used to reduce the Bravyi-

Kitaev transformed Hamiltonian to a 2-local programmable system with a restricted set of physical couplings. Finally, tunneling spectroscopy of a probe qubit [29] can be used to measure eigenvalues of the prepared state directly.

While the exact length of time one must adiabatically evolve is generally unknown, Babbush *et al.* argue that the excited state gap could shrink polynomially with the number of spin-orbitals when interpolating between exactly preparable noninteracting subsystems and the exact molecular Hamiltonian in which those subsystems interact. This would imply that adiabatic state preparation is efficient. Their argument is based on the observation that molecular systems are typically stable in their electronic ground states and the natural processes which produce these states should be efficient to simulate with a quantum device. Subsequently, Veis and Pittner analyzed adiabatic state preparation for a set of small chemical systems and observed that for all configurations of these systems, the minimum gap occurs at the very end of the evolution when the state preparation is initialized in an eigenstate given by a CAS (complete active space) ground state [272]. The notion that the minimum gap could be bounded by the physical HOMO (highest occupied molecular orbital) - LUMO (lowest unoccupied molecular orbital) gap lends support to the hypothesis put forward by Babbush *et al.*

### 9.4.1 Resources for adiabatic quantum chemistry

In the adiabatic model of quantum computation, the structure of the final problem Hamiltonian (encoding the molecular eigenspectrum) determines experimental resource requirements. Since programmable many-body interactions are generally unavailable, we will assume that any experimentally viable problem Hamiltonian must be 2-local. Any 2-local

Hamiltonian on $n$ qubits can be expressed as,

$$H = \alpha \cdot \mathbf{1} + \sum_{i=1}^{n} \vec{\beta}_i \cdot \vec{\sigma}_i + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \vec{\gamma}_{ij} \cdot (\vec{\sigma}_i \otimes \vec{\sigma}_j) \tag{9.61}$$

where $\vec{\sigma}_i = \langle \sigma_i^x, \sigma_i^y, \sigma_i^z \rangle$ is the vector of Pauli matrices on the $i^{\text{th}}$ qubit, $\alpha \in \mathbb{R}$ is a scalar and $\vec{\beta}_i \in \mathbb{R}^3$ and $\vec{\gamma}_{ij} \in \mathbb{R}^9$ are vectors of coefficients for each possible term.

In addition to the number of qubits, the most important resources are the number of qubit couplings and the range of field values needed to accurately implement the Hamiltonian. Since local fields are relatively straightforward to implement, we are concerned with the number of 2-local couplings,

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \text{card}\left(\vec{\gamma}_{ij}\right) \tag{9.62}$$

where $\text{card}\left(\vec{v}\right)$ is the number of nonzero terms in vector $\vec{v}$. Since the effective molecular electronic structure Hamiltonian is realized perturbatively, there is a tradeoff between the error in the eigenspectrum of the effective Hamiltonian, $\epsilon$, and the strength of couplings that must be implemented experimentally. The magnitude of the perturbation is inversely related to the gadget spectral gap $\Delta$ which is directly proportional to the largest term in the Hamiltonian,

$$\max_{ij} \left\{ \left\| \vec{\gamma_{ij}}\left(\epsilon\right) \right\|_{\infty} \right\} \propto \Delta\left(\epsilon\right). \tag{9.63}$$

Thus, the smaller $\Delta$ is, the easier the Hamiltonian is to implement but the greater the error in the effective Hamiltonian. In general, there are other important resource considerations but these are typically scale invariant; for instance, the geometric locality of a graph or the set of allowed interaction terms. The Hamiltonian can be modified to fit such constraints using additional perturbative gadgets but typically at the cost of using more ancilla qubits that require greater coupling strength magnitudes.

### 9.4.2 Estimates of qubit and coupler scaling

The Bravyi-Kitaev transformation is crucial when embedding molecular electronic structure in 2-local spin Hamiltonians due to the fact that this approach guarantees a logarithmic upper-bound on the locality of the Hamiltonian. A loose upper-bound (i.e. overestimation) for the number of qubits needed to gadgetize the molecular electronic Hamiltonian can be obtained by assuming that all terms have the maximum possible locality of $O(\log(M))$ where $M$ is the number of spin-orbitals.

In general, the number of terms produced by the Bravyi-Kitaev transformation scales the same as the number of integrals in the electronic structure problem, $O\left(M^4\right)$; however, as pointed out in an earlier section, this bound can be reduced to $O\left(M^2\right)$ if a local basis is used and small integrals are truncated. Using the "bit-flip" gadgets of [165, 157] to reduce $M^2$ terms of locality $\log(M)$, we would need $M^2 \log(M)$ ancillae. Since the number of ancilla qubits is always more than the number of logical qubits for this problem, an upper-bound on the total number of qubits needed is $O\left(M^2 \log(M)\right)$.

The number of couplings needed will be dominated by the number of edges introduced by ancilla systems required as penalty terms by the bit-flip gadgets. Each of the $O\left(M^2\right)$ terms is associated with a different ancilla system which contains a number of qubits equal to the locality of that term. Furthermore, all qubits within an ancilla system are fully connected. Thus, if we again assume that all terms have maximum locality, an upper-bound on the number of couplers is $O\left(M^2 \log^2(M)\right)$. Based on this analysis, the adiabatic approach to quantum chemistry has rather modest qubit and coupler requirements.

### 9.4.3 Estimates of spectral gap scaling

In [14], Babbush *et al.* reduce the locality of interaction terms using perturbative gadgets from the "bit-flip" family, first introduced in [165] and later generalized by [157]. In the supplementary material presented in a later paper analyzing the scaling of gadget constructions [63], it is shown that for bit-flip gadgets, $\lambda^{k+1}/\Delta^k = O(\epsilon)$ and

$$\max_{ij} \left\{ \left\| \vec{\gamma_{ij}}(\epsilon) \right\|_\infty \right\} = O\left( \frac{\lambda^k}{\Delta^{k-1}} \right). \tag{9.64}$$

Here, $\lambda$ is the perturbative parameter, $\Delta$ is the spectral gap, $\epsilon$ is the error in the eigenspectrum and $\vec{\gamma}_{ij}$ is the coefficient of the term to be reduced. Putting this together and representing the largest coupler value as $\gamma$, we find that $\Delta = \Omega\left(\epsilon^{-k}\gamma^k\right)$, where $\Omega$ is the "Big Omega" lower bound. Due to the Bravyi-Kitaev transformation, the locality of terms is bounded by, $k = O(\log(M))$; thus, $\Delta = \Omega\left(\epsilon^{-\log(M)}\gamma^{\log(M)}\right)$.

Prior analysis from this paper indicates that the maximum integral size is bounded by $\gamma \leq |\beta_{\max}^{\mathrm{OEI}}| M^{2/3}$. This gives us the bound,

$$\Delta = \Omega\left( \epsilon^{-\log(M)} \left\| \beta_{\max}^{\mathrm{OEI}} M^{2/3} \right\|^{\log(M)} \right). \tag{9.65}$$

However, $\Delta$ also depends polynomially on $M^2$, the number of terms present. Though known to be polynomial, it is extremely difficult to predict exactly how $\Delta$ depends on $M^2$ as applying gadgets to terms "in parallel" leads to "cross-gadget contamination" which contributes at high orders in the perturbative expansion of the self-energy used to analyze these gadgets [63]. Without a significantly deeper analysis, we can only conclude that,

$$\Delta = \Omega\left( \mathrm{poly}(M) \left\| \frac{\beta_{\max}^{\mathrm{OEI}} M^{2/3}}{\epsilon} \right\|^{\log(M)} \right). \tag{9.66}$$

This analysis indicates that the most significant challenge to implementing the adiabatic approach to quantum chemistry is the required range of coupler values which is certain to span *at least* several orders of magnitude for non-trivial systems.

This calls attention to an important open question in the field of Hamiltonian gadgets: whether there exist "exact" gadgets which can embed the ground state energy of arbitrary many-body target Hamiltonians without the use of perturbation theory. A positive answer to this conjecture would allow us to embed molecular electronic structure Hamiltonians without needing large spectral gaps. For entirely diagonal Hamiltonians, such gadgets are well known in the literature [16] but fail when terms do not commute [63]. Exact reductions have also been achieved for certain Hamiltonians. For instance, "frustration-free" gadgets have been used in proofs of the QMA-Completeness of quantum satisfiability, and in restricting the necessary terms for embedding quantum circuits in Local Hamiltonian [203, 123, 70].

## 9.5 Conclusions

In this work, we analyzed the impact on scaling for quantum chemistry on a quantum computer that results from consideration of locality of interactions and exploitation of local basis sets. The impact of locality has been exploited to great advantage for some time in traditional algorithms for quantum chemistry, but has received relatively little attention in quantum computation thus far. From these considerations, we showed that in practical implementations of quantum phase estimation using a first order Suzuki-Trotter approximation, one expects a scaling cost on the order of $O(M^5 \log M)$ with respect to number of spin-orbitals, rather than more pessimistic estimates of $O(M^8)$-$O(M^9)$[281, 140] or $O(M^{5.5})$-$O(M^{6.5})$[224] related to the use of unphysical random integral distributions or the restriction to molecules too small to observe the effects of physical locality. We believe that the combination of the algorithmic improvements suggested by Poulin and Hastings et al [224, 140] with strategies that exploit locality presented here will result in even greater

gains, and more work is needed in this area.

We also considered the cost of Hamiltonian averaging, an alternative to quantum phase estimation with minimal coherence time requirements beyond state preparation. This method has some overhead with respect to quantum phase estimation, scaling as $O(M^6)$ in the number of spin-orbitals, but has significant practical advantages in coherence time costs, as well as the ability to make all measurements in parallel. This method can at best give the energy of the state provided when oracle guesses are imperfect, however it can easily be combined with a variational or adiabatic approach to improve the accuracy of the energy estimate. Moreover, while quantum phase estimation promises to be able to remove the bias of imperfect oracle guesses, we demonstrated how the cost of removal may strongly depend on how imperfect the guesses are.

Finally we analyzed the impact of locality on a complementary approach for quantum chemistry, namely adiabatic quantum computation. This approach does not have a known direct dependence on the quality of guess states provided by an oracle, and can in fact act as the state oracle for the other approaches discussed here.

In all cases, the consideration of physical locality greatly improves the outlook for quantum chemistry on a quantum computer, and in light of the goal of quantum chemistry to study physical systems rather than abstract constructs, it is the correct to include this physical locality in any analysis pertaining to it. We believe that with these and other developments made in the area of quantum computation, quantum chemistry remains one of the most promising applications for exceeding the capabilities of current classical computers.

## Acknowledgments

J.M. is supported by the Department of Energy Computational Science Graduate Fellowship under grant number DE-FG02-97ER25308. P.J.L. acknowledges the National Science Foundation grant number PHY-0955518. A.A-G. and P.J.L. appreciate the support of the Air Force Office of Scientific Research under Award No. FA9550-12-1-0046. R.B. and A.A.-G. are grateful to the National Science Foundation for Award No. CHE-1152291.

## Chapter 10

# Quantum Chemistry Simulation in a Solid-State Spin Register

Apart from minor modifications, this chapter originally appeared as [279]:

"Quantum Simulation of Helium Hydride in a Solid-State Spin Register". Ya Wang, Florian Dolde, Jacob Biamonte, Ryan Babbush, Ville Bergholm, Sen Yang, Ingmar Jakobi, Philipp Neumann, Alán Aspuru-Guzik, James D. Whitfield and Jørg Wrachtrup. arXiv preprint 1405.2696. 1-9. 2014.

## Abstract

*Ab initio* computation of molecular properties is one of the most promising applications of quantum computing. While this problem is widely believed to be intractable for classical computers, efficient quantum algorithms exist which have the potential to vastly accelerate research throughput in fields ranging from material science to drug discovery. Using a solid-state quantum register realized in a nitrogen-vacancy (NV) defect in diamond, we compute the bond dissociation curve of the minimal basis helium hydride cation, $HeH^+$. Moreover, we report an energy uncertainty (given our model basis) of the order of $10^{-14}$ Hartree, which is ten orders of magnitude below desired chemical precision. As NV centers in diamond provide a robust and straightforward platform for quantum information processing, our

work provides several important steps towards a fully scalable solid state implementation of a quantum chemistry simulator.

## 10.1 Introduction

Quantum simulation, as proposed by Feynman [107] and elaborated by Lloyd [183] and many others [289, 298, 1, 33, 117], exploits the inherent behavior of one quantum system as a resource to simulate another quantum system. Indeed, there have been several experimental demonstrations of quantum simulators in various architectures including quantum optics, trapped ions, and ultracold atoms [265]. The importance of quantum simulators applied to electronic structure problems has been detailed in several recent review articles including [297, 163, 58, 186, 116] and promises a revolution in areas such as materials engineering, drug design and the elucidation of biochemical processes.

The computational cost of solving the full Schrödinger equation of molecular systems using any known method on a classical computer scales exponentially with the number of atoms involved. However, there is now a growing body of work proposing efficient quantum simulations of chemical Hamiltonians, e.g. [11, 175, 285, 163, 77, 281, 264, 140]. A general procedure to obtain molecular eigenenergies to a desired precision is: $(i)$ mapping molecular wave functions into the computational basis, $(ii)$ preparing the quantum simulator into an ansatz state which is close to an eigenstate of the simulated Hamiltonian $H_{\rm sim}$, $(iii)$ encoding the energies into a relative phase by simulating the time evolution operator $e^{-itH_{\rm sim}}$ using quantum gates, and $(iv)$ extracting the energies to desired precision using a variant of the quantum phase estimation algorithm [11, 183, 169]. Experimental realizations of the quantum simulation of electronic structure began with the simulation of molecular hydrogen using quantum optics [175] and liquid state NMR [97]. A calculation of the energy of the

helium hydride cation in a photonics setup using a quantum variational eigensolver that avoids phase estimation has also been performed [223].

Nitrogen-vacancy (NV) centers in diamond offer a scalable and precise platform for quantum simulation which does not suffer from signal losses as the system size increases and can avoid challenges such as the need for post-selected measurements. Progress to date has shown that such systems are among the most accurate and most controllable candidates for quantum information processing [125, 206, 230, 275, 207, 263, 96, 31, 95, 260, 22, 193]. Milestone demonstrations include high-fidelity initialization and readout [125, 206, 230, 275], on-demand generation of entanglement [207, 263, 96, 31, 275, 95], implementation of quantum control [151, 99, 95], ultra-long spin coherence time [22], non-volatile memory [193], quantum error correction [275, 260], as well as a host of metrology and sensing experiments [261, 94]. Several proposals to scale up the size of NV systems currently exist, e.g. [30, 95]. Building on this premise, this is the first study reporting the use of a solid state spin system to simulate quantum chemistry.

The chemical system we consider in this paper is the helium hydride cation, $HeH^+$ (see Fig. 10.1a), believed to be the first molecule in the early universe [101]. While $HeH^+$ is isoelectronic (i.e. has the same number of electrons) with the previously studied molecular hydrogen, the reduced symmetry requires that we simulate larger subspaces of the full configuration interaction (FCI) Hamiltonian $H_{\text{sim}}$. Specifically, we consider

$$H_{\text{sim}} = T_e + W_{ee} + V_{eN}(R) + E_N(R) \tag{10.1}$$

in a minimal single particle basis with one site per atom. Here, $T_e$ and $W_{ee}$ are the kinetic and Coulomb operators for the electrons, $V_{eN}$ is the electron-nuclear interaction, and $E_N$ is the nuclear energy due to the Coulomb interaction between the hydrogen and helium atoms. The last two terms depend on the internuclear distance $R$.

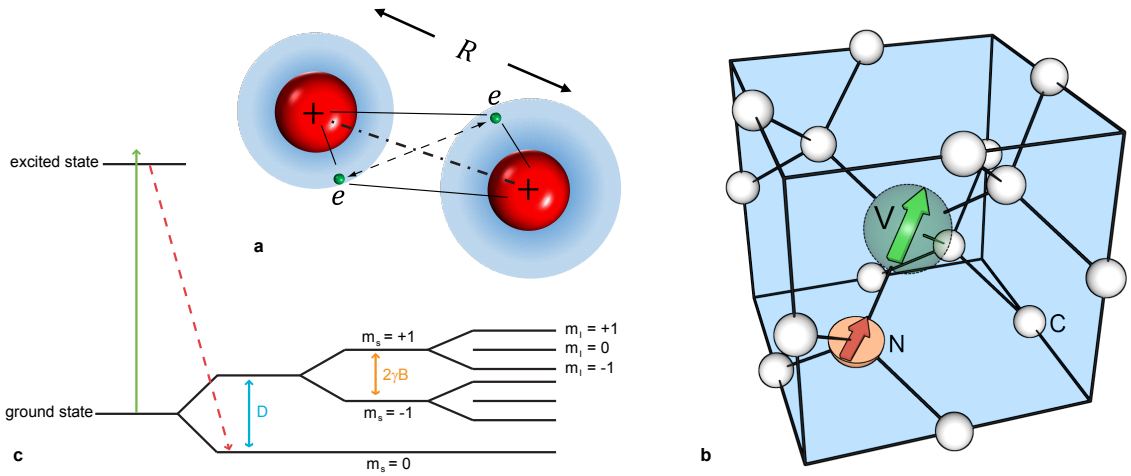



Figure 10.1: **Calculation of $HeH^+$ molecular energy with NV spin register in diamond** (a) $HeH^+$, molecule to be simulated. It consists of a hydrogen and a helium nucleus, and two electrons. The distance (bond length) between the nuclei is denoted by $R$. Dot-dashed line, straight line, and dotted arrow indicate the nucleus-nucleus, electron-nucleus and electron-electron Coulomb interactions, respectively. (b) A nitrogen-vacancy center in diamond, used as a quantum simulator. The electron spin is used for simulation and the nuclear spin as the probe qubit for energy readout. (c) Energy level diagram for the coupled spin system formed by the NV electron spin and nearby $^{14}$N nuclear spin. Optical transitions between ground and excited state are used to initialize and measure the electron spin state.

In this work, we consider the singlet ($S = 0$) sector of the electronic Hamiltonian in a minimal single-electron basis consisting of a single site at each atom given by contracted Gaussian orbitals. After taking symmetries into account, the Hamiltonian can be represented as a $3 \times 3$ matrix in the basis $(\Psi_1,\ \Psi_6,\ \frac{1}{\sqrt{2}}(\Psi_3 - \Psi_4))$ (see Methods). Each term of the Hamiltonian in the single particle basis (e.g. $\langle\chi_i|(T_e + V_{eN})|\chi_j\rangle$) is precomputed classically at each internuclear separation $R$ using the canonical spin orbitals found via the Hartree-Fock (HF) procedure.

After obtaining $H_{\text{sim}}$ through this (typically) efficient classical computation, we perform the quantum simulation of this molecule on a single-NV register, which consists of an electronic spin-1 and an associated $^{14}$N nuclear spin-1 forming a qutrit pair (see Fig. 10.1b).

The electronic spin-1 of the NV system acts as the *simulation register* through mapping the molecular basis ($\Psi_1$, $\Psi_6$, $\frac{1}{\sqrt{2}}(\Psi_3 - \Psi_4)$) onto its $m_s = (1, 0, -1)$ states. The $^{14}$N nuclear spin-1 is used as the *probe register* to read out the energies using the iterative phase estimation algorithm (IPEA) [219], as shown in Fig. 10.1c. The controlled evolution $e^{-itH_{\text{sim}}}$ on the electron spin is implemented using optimal control theory, which helps to realize the most precise simulation of quantum chemistry to date. Without post-selection and at room temperature, our experimentally computed energy agrees with the corresponding classical calculations to within chemical precision, with a deviation of $10^{-14}$ Hartree. By performing the simulation process for different values of $R$, the electronic potential energy surfaces are also experimentally obtained.

In order to efficiently sample the eigenenergy $E_n$ as the size of the system grows, one must prepare an ansatz state that has an overlap with the corresponding eigenstate $|e_n\rangle$ that decreases at most polynomially in the system size. The phase estimation algorithm [169] can then be used to project the ansatz state into the exact eigenstate with sufficiently high probability.

One possible approach to realize this requirement is to use adiabatic state preparation [11, 97, 14, 272], the performance of which depends on the energy gap during the entire evolution process. An alternative approach is to approximate the eigenstate with a trial state. Such trial states can often be prepared based on classical approximate methods. In our case, the simulation register is initialized in a trial state $|\tau\rangle \in \{|+1\rangle, |-1\rangle\}$, expressible as a superposition of all the $H_{\text{sim}}$ eigenstates, $|\tau\rangle = \sum_k a_k |e_k\rangle$. The probe register is prepared in the state $|\psi(0)\rangle = (|0\rangle + |-1\rangle)/\sqrt{2}$ (see Methods).

In the next step, a controlled-$U(t)$ gate for different times $t$, where $U(t) = \exp(-iH_{\text{sim}}t)$,

is applied to encode the energies into a relative phase, resulting in the state

$$|\psi(t)\rangle = \frac{1}{\sqrt{2}} \sum_k a_k (|0\rangle + e^{-iE_k t} |-1\rangle) |e_k\rangle \,. \tag{10.2}$$

The reduced density matrix of the probe register,

$$\rho_{\text{probe}}(t) = \frac{1}{2} \begin{pmatrix} 1 & \sum_k |a_k|^2 e^{-iE_k t} \\ \sum_k |a_k|^2 e^{iE_k t} & 1 \end{pmatrix}, \tag{10.3}$$

contains the information about the energies in its off-diagonal elements. This information is then transferred to the electron spin for readout by a nuclear spin $\frac{\pi}{2}$-pulse and selective $\pi$-pulses on the electron spin-1 (Fig. 10.2a).

To measure the energy precisely, we perform classical Fourier analysis on the signal for different times $(t_s, 2t_s, \ldots, Lt_s)$. This readout method can help to resolve the probability $|a_k|^2$ of each eigenstate $|e_k\rangle$ and approximate the corresponding energy $E_k$. We choose $t_s$ such that the sampling rate $\frac{1}{t_s} > |E_n|/\pi$. To enhance the precision of the energy eigenvalues, an iterative phase estimation algorithm is performed. A central feature of this algorithm includes repeating the unitary operator $U$ to increase readout precision. Expressing the energy as a string of decimal digits, $E_k = x_1.x_2x_3\ldots$, the first digit $x_1$ can be determined by the first round phase estimation process. Once $x_1$ is known, the second digit $x_2$ can be iteratively determined by implementing the unitary operator $U^p$, where $p = 10$. For the $k^{\text{th}}$ iteration, $p = 10^{k-1}$.

An increasingly precise energy can be obtained through continued iterations. However, the repetitions and therefore the iterations are fundamentally limited by the coherence time of the quantum system. Moreover, the accumulated gate errors become a dominant limitation of the energy precision as the repetitions increase. To avoid such shortcomings, the time evolution operators $U^p$ are realized and optimized with optimal control theory (see Methods). The precision we reach in our experiments demonstrates that optimal control can

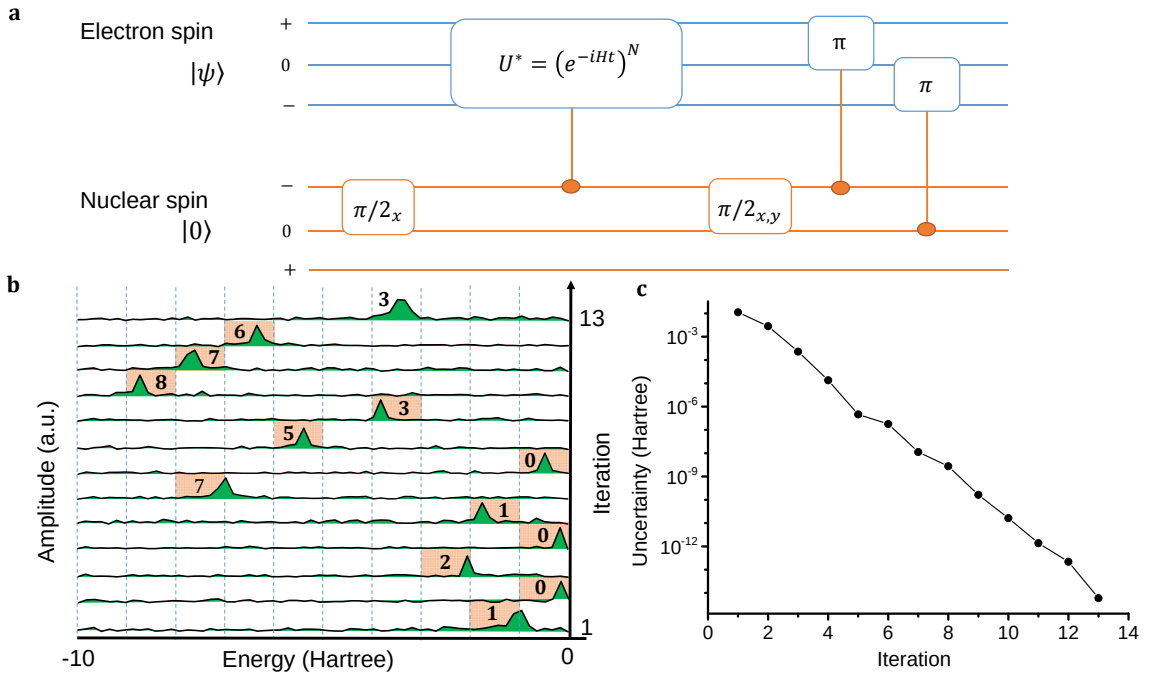



Figure 10.2: **Energy readout through quantum phase estimation algorithm** (a) Experimental implementation of the IPEA algorithm. The controlled gate $U^*$ is realized using optimal control (see Methods). The $x$, $y$ phases in the last $\pi/2$ pulse measure the real and imaginary parts of the signal, respectively, which yield the sign of the measured energy. The number of repetitions $N = 10^{k-1}$ depends on the iteration $k$. (b) Experimental results of iterative phase estimation algorithm to enhance the precision of measured energy for the case of $R = 0.9$. The Fourier spectrum of the first iteration ($k = 1$) fixes the energy roughly between $-10$ and 0 Hartree. The precision is then improved iteratively by narrowing down the energy range. In each iteration, the energy range is divided into ten equal segments. The red area indicates the energy range for the next iteration. After each iteration at least one decimal digit, denoted by the number in the red area, is resolved. (c) The uncertainty of the measured energy as a function of the iteration number.

overcome several difficult features found when scaling up the register size [95]. Although it cannot be applied in large registers to generate the quantum gates directly, it can be used to generate flexible smaller building blocks, ensuring high-fidelity control in future large scale applications. In the present case, the method is unscalable because we compute the unitary propagator using a classical computer. However, by using a Trotter-type gate sequence to implement the propagators, e.g. [285], this can be designed with polynomially scaling.

Fig. 10.2b shows our results of internuclear distance $R = 90$ pm with trial state $|+1\rangle$. The position of the peak indicates the eigenvalue of molecular Hamiltonian with an offset $\mathrm{tr}(H_{\mathrm{sim}})/3$. The Fourier spectrum has only one major peak, which shows that the trial state $|+1\rangle$ is close to the ground state. As the iterations increase, more precise decimal digits of the ground state energy are resolved. After 13 repetitions the molecular energy is extracted to be $-1.020170538763387 \pm 8 \times 10^{-15}$ Hartree, very close to the theoretic value, which is -1.020170538763381 Hartree, with an uncertainty of $\pm 1.4 \times 10^{-14}$ Hartree.

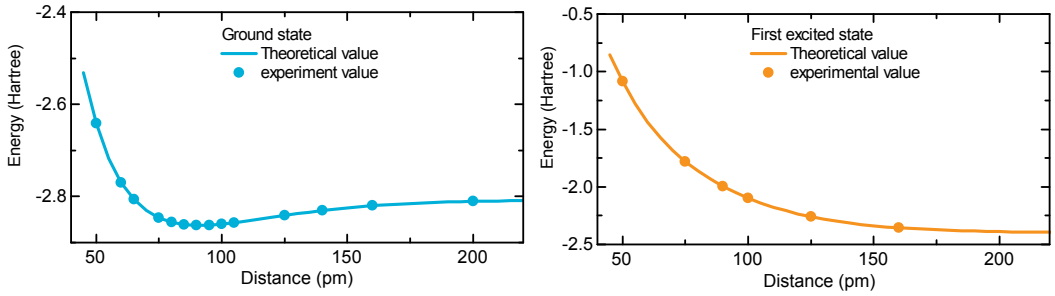



Figure 10.3: **Energy surfaces of the $HeH^+$ molecule.** The energy surface of the second excited state can be obtained by subtracting energies of the the ground and first excited states from the trace of $H_{\mathrm{sim}}$, and is not shown. All the measured energies are obtained in five iterations.

Once the energies have been measured, we can obtain the potential energy surface of the molecule by repeating the procedure for different distances $R$ (see Fig. 10.3). The ground state energy surface is obtained with trial state $|+1\rangle$ and first excited state energy surface is obtained with trial state $|-1\rangle$.

We obtain the remaining eigenenergy (of the second excited state) without further measurement by subtracting the ground and first excited state energies from the trace of $H_{\mathrm{sim}}$. The potential energy surfaces can be used to compute key molecular properties such as ionization energies and vibrational energy levels. An important example is the equilibrium geometry: we found the minimal energy for the ground state, $-2.86269$ Hartree,

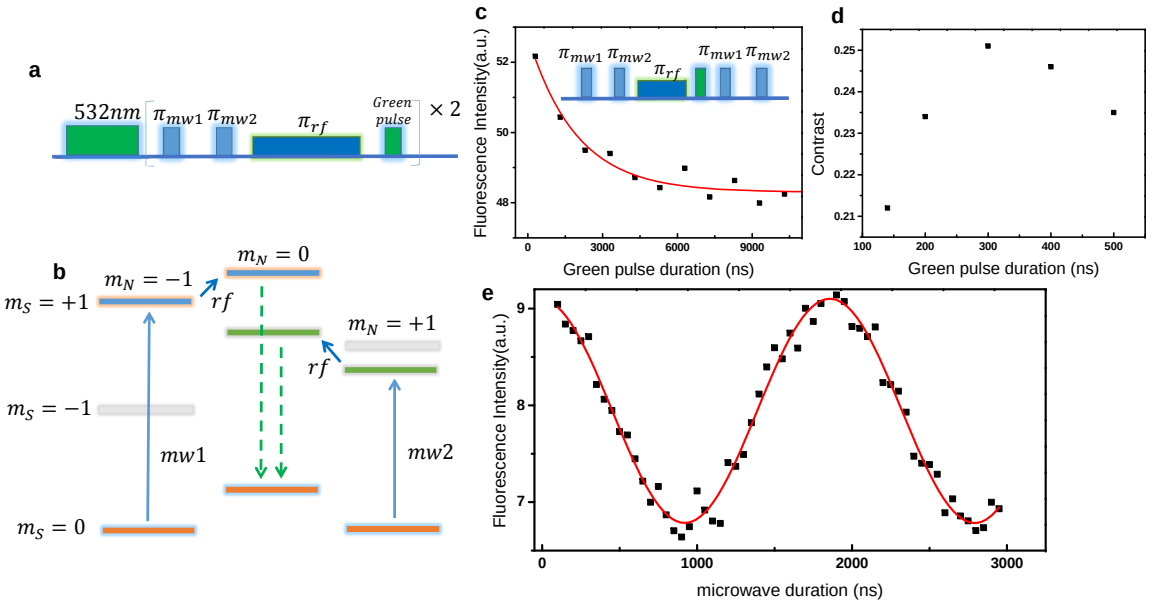



Figure 10.4: **Polarization of $^{14}$N nuclear spin.** (a) Experimental pulse sequence to polarize the $^{14}$N nuclear spin. (b) Dynamical process during one polarization step. (c) Initialization decay of nuclear spin under laser illumination. The fit (red line) shows an exponential decay with time constant $1.9 \pm 0.3$ $\mu$s. The pulse sequence is shown in the inset. (d) The contrast of electron spin Rabi oscillation in the $m_N = 0$ subspace (here: a measure for the degree of nuclear spin initialization) varies with the second laser pulse length. (e) The electron spin Rabi oscillation in the $m_N = 0$ subspace for 300 ns green pulse duration.

at a bond length of 91.3 pm. In addition, we obtained a binding energy of 0.07738 Hartree in our basis. To improve the accuracy of our results we would need to simulate the system in a larger basis, thereby requiring more qutrits.

## 10.2 Methods

### 10.2.1 Computation of molecular Hamiltonians

The full configuration interaction Hamiltonian is a sparse matrix and each matrix element can be computed in polynomial time. For a basis set with $M$ orbitals, there are $M^4$ terms in the Hamiltonian but the Hamiltonian is of size $\frac{M!}{N!(M-N)!} \approx M^N$ which is

exponential as the number of electrons grow. To generate the Hamiltonian, we fix the nuclear configuration and then compute the necessary one- and two-body integrals which parameterize the FCI matrix at each fixed bond length in the standard STO-3G basis [144], using the PSI3 electronic structure package [80].

The minimal basis $HeH^+$ system has two spatial orbitals which we denote as $g(r)$ and $e(r)$ and two spin functions denoted as $\alpha(\sigma)$ and $\beta(\sigma)$ which are eigenstates of the $S_z$ operator. We combine these to form four spin orbitals, $\chi_1 = g(r)\alpha(r)$, $\chi_2 = g(r)\beta(\sigma)$, $\chi_3 = e(r)\alpha(\sigma)$ and $\chi_4 = e(r)\beta(\sigma)$. There are six possible two-electron Slater determinants, $\Psi_1 = \mathcal{A}(\chi_1\chi_2)$, $\Psi_2 = \mathcal{A}(\chi_1\chi_3)$, $\Psi_3 = \mathcal{A}(\chi_1\chi_4)$, $\Psi_4 = \mathcal{A}(\chi_2\chi_3)$, $\Psi_5 = \mathcal{A}(\chi_2\chi_4)$, and $\Psi_6 = \mathcal{A}(\chi_3\chi_4)$. More explicitly,

$$\mathcal{A}(\chi_i\chi_j) = \frac{1}{\sqrt{2}} \begin{vmatrix} \chi_i(r_1\sigma_1) & \chi_j(r_1\sigma_1) \\ \chi_i(r_2\sigma_2) & \chi_j(r_2\sigma_2) \end{vmatrix}. \tag{10.4}$$

States $\Psi_1$, $\Psi_3$, $\Psi_4$, and $\Psi_6$ have total projected spin of $M_z = 0$ whereas $\Psi_2$ and $\Psi_5$ have projected values of $M_z = 1$ and $M_z = -1$ respectively. Only $\Psi_1$ and $\Psi_6$ are valid eigenstates of the total spin operator $S^2$; however, the symmetric and antisymmetric combinations of $\Psi_3$ and $\Psi_4$ yield the $m_s = 0$ triplet and an additional singlet, respectively. When a computation is requested on the singlet state, the PSI3 package computes the symmetry-adapted FCI matrix in the basis of $\Psi_1$, $\Psi_3$, $\Psi_4$ and $\Psi_6$. By combining $\Psi_3$ and $\Psi_4$ we obtained the three $HeH^+$ singlet states used in this experiment: $\Psi_1$, $\Psi_6$ and $\frac{1}{\sqrt{2}}(\Psi_3 - \Psi_4)$.

### 10.2.2 Sample characteristics

We use a nitrogen-vacancy center in high-purity diamond grown by microwave-assisted chemical vapor deposition (CVD). The intrinsic nitrogen content of the grown crystal is below 1 ppb and the $^{12}C$ content is enriched to 99.9%. Experiments are performed at room

temperature with an applied magnetic field of 11 gauss. The electron spin's coherence times are $T_2^* \approx 80$ $\mu$s and $T_2 \approx 600$ $\mu$s.

### 10.2.3 NV system

In a magnetic field $B_0$ aligned along the NV symmetry axis, the electronic and nuclear spin system has the Hamiltonian

$$H/\hbar = 2\pi\Delta S_z^2 + \gamma_e B_0 S_z + 2\pi A_{\rm hf} S_z I_z + 2\pi Q I_z^2 + \gamma_N B_0 I_z$$

where $S_z$ and $I_z$ are the dimensionless spin-1 operators for the electrons and the $^{14}$N nucleus, respectively. $\Delta \approx 2.87$ GHz and $Q \approx -4.94$ MHz are the zero-field splitting of the electronic spin and quadrupole splitting of the nuclear spin. The hyperfine coupling coefficient is $A_{\rm hf} \approx 2.16$ MHz. The Larmor frequencies are defined as $\omega_i := \gamma_i B_0$, where $\gamma_i$ is the gyromagnetic ratio of the spin (electronic or nuclear).

### 10.2.4 System initialization

In the experiment, the $^{14}$N nuclear spin is initially in a thermal state. It is polarized into the spin state $|m_I = 0\rangle$ by means of optical pumping of the electron spin followed polarization transfer realized with electron spin and nuclear spin control (see Fig. 10.4). The second short laser pulse repolarizes the electron spin into $|m_s = 0\rangle$, leaving the spins in the state $|m_s = 0, m_I = 0\rangle$ [99]. In practice, the imperfect control and short $T_1 \approx 1.9$ $\mu$s time of nuclear spin under laser illumination will result in imperfect polarization of the nuclear spin. To enhance the polarization effect, we repeat the process two times and tune the second laser pulse to an optimal length around 300 ns. The observed electron spin Rabi oscillation in the $m_I = 0$ subspace indicates a final polarization of around 60%. After the polarization process, the electron spin is then prepared into the $|m_s = +1\rangle$ or $|m_s = -1\rangle$

state by another microwave $\pi$ pulse unconditional on the nuclear spin state. Note that only the phase of the nuclear spin superposition state contains information in the IPEA process, therefore imperfect polarization would not affect the accuracy of final energy measurement.

### 10.2.5 Controlled U(t) gate realization

In the experiment, every individual controlled gate $U^{'} = (e^{-iH_{\rm sim}t})^p$ can be realized by decomposing it into more basic but highly complicated microwave pulses. However, this approach will accumulate considerable control errors. To avoid such shortcomings, we use an alternative method, optimal control, which has recently been used to achieve high-fidelity control in coupled NV centers in diamond [95].

To make the calculation feasible, another equivalent controlled gate $U^* = e^{-iH't}$ with the Hamiltonian $H' = H_{\rm sim} - {\rm tr}(H_{\rm sim})/3$ is calculated. This operation will only introduce additional $O(1)$ complexity. One then needs to add this constant value ${\rm tr}(H_{\rm sim})/3$ back to the final measured energies. To calculate $U^*$, we use the GRAPE algorithm [166] to optimize the pulse sequence, with the final fidelity always larger than 0.99. For every controlled gate, the pulse sequence consists of 10 pieces of 140 ns each. Two microwave frequencies are applied simultaneously to control the electron spin, in the observed hyperfine peaks of the $|m_I = -1, m_s = 0\rangle \rightarrow |m_I = -1, m_s = +1\rangle$ and $|m_I = -1, m_s = 0\rangle \rightarrow |m_I = -1, m_s = -1\rangle$ transitions. More details about the optimal control method can be found in reference [95].

### 10.2.6 A symmetry of the ground state energy problem

If we write the system Hamiltonian as $H = T + K$ where diagonal $T$ accounts for the HF approximations and off-diagonal $K$ accounts from the Born-Oppenheimer approximate treatment of the problem. We note that whenever the support of $K$ corresponds to the

adjacency matrix of a bipartite graph, then $H = T + K$ and $L = T - K$ are cospectral. This follows from the proof [299] that any bipartite (necessarily time-inversion symmetric) Hamiltonian $H$ is on the same orbit as $-H$ under conjugation by diagonal unitarians (e.g. there exists a diagonal unitary $\Lambda$ such that $\Lambda H \Lambda^\dagger = -H$) where $T$ is central under this action. Hence, they represent the same ground-state energy problems, providing an equivalent problem instance $L$ to attempt state preparation on. It turns out that all of the quantum chemistry algorithms realized to date [175, 97, 187] have this property including our own demonstration, where the underlying graph corresponds to a tree. This observation provides a second benchmark to be considered in future experiments.

## 10.3 Discussion

We will now briefly discuss several of the implications of this study. Current quantum simulations cannot outperform classical devices. In large systems, the simulated propagators can be implemented using Trotter sequences and should be accompanied by error correction. Optimal control methods, as we have demonstrated here, should prove necessary to perform these tasks with satisfactory precision.

We have demonstrated the most precise quantum simulation of molecular energies to date, which represents an important step towards the advanced level of control required by future quantum simulators that will outperform classical methods. The energies we obtained for the helium hydride cation surpass chemical precision by 10 orders of magnitude (with respect to the basis). The accuracy of our results can be increased by using a larger, more flexible single-particle basis set but this will require a larger quantum simulator that eventually will require error correction schemes [77].

Our study presents evidence that quantum simulators can be controlled well enough

to recover increasingly precise data. The availability of highly accurate energy eigenvalues of large molecules is presently far out of reach of existing computational technology, and quantum simulation could open the door to a vast range of new technological applications. The approach we took was based on iterative phase estimation [219] and optimal control decompositions [95]—these will form key building blocks for any solid-state quantum simulator. Even more generally, this study would suggest that the techniques presented here should be employed in any future simulator that will outperform classical simulations of electronic structure calculations.

## Acknowledgements

V.B. and J.D.B. acknowledge financial support by Fondazione Compagnia di San Paolo through the Q-ARACNE project. J.D.B. would also like to acknowledge the Foundational Questions Institute (under grant FQXi-RFP3-1322) for financial support. R.B. and A.A.-G. acknowledge support from the Air Force Office of Scientific Research under contract FA9550-12-1-0046, as well as the National Science Foundation CHE-1152291 and the Corning Foundation. J.W. acknowledges support by the EU via IP SIQS and the ERC grant SQUTEC as well as the DFG via the research group 1493 and SFB/TR21 and the Max Planck Society. J.D.W. thanks the VCQ and Ford postdoctoral fellowships for support. We thank Mauro Faccin and Jacob Turner for providing valuable feedback regarding the manuscript.

# Bibliography


[1] D. S. Abrams and S. Lloyd. Simulation of Many-Body Fermi Systems on a Universal Quantum Computer. *Phys. Rev. Lett.*, **79**, 4 (1997).

[2] D. S. Abrams and S. Lloyd. Quantum Algorithm Providing Exponential Speed Increase for Finding Eigenvalues and Eigenvectors. *Phys. Rev. Lett.*, **83**, 5162 (1999).

[3] D. Aharonov, J. Kempe, S. Lloyd, W. V. Dam, Z. Landau, O. Regev, W. Van Dam, J. Kempe, Z. Landau, S. Lloyd, and O. Regev. Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM J. Comput.*, **37**, 166 (2007).

[4] D. Aharonov and A. Ta-Shma. Adiabatic quantum state generation and statistical zero knowledge. In *Proc. thirty-fifth ACM Symp. Theory Comput. - STOC '03*, page 20, New York, New York, USA, June 2003. ACM Press.

[5] F. A. Aloul, A. Ramani, I. L. Markov, and K. A. Sakallah. Generic ILP versus specialized 0-1 ILP: an update. In *IEEE/ACM Int. Conf. Comput. Aided Des. 2002. ICCAD 2002.*, pages 450–457. Ieee, 2002.

[6] M. Amin. Searching for quantum speedup in quasistatic quantum annealers. *e-print arXiv: 1503.04216*, page 6 (2015).

[7] A. Apolloni, N. Cesa-Bianchi, and D. De Falco. A numerical implementation of quantum annealing. In *Stoch. Process. Phys. Geom. Proc. Ascona-Locarno Conf.*, pages 97–111, 1988.

[8] B. Apolloni, C. Carvalho, and D. De Falco. Quantum stochastic optimization. *Stoc. Proc. Appl.*, **33**, 233 (1989).

[9] S. Arora, C. Lund, R. Motwani, M. Sudan, and M. Szegedy. Proof verification and the hardness of approximation problems. *J. ACM*, **45**, 501 (1998).

[10] E. Artacho, D. Sánchez-Portal, P. Ordejón, A. Garcia, and J. M. Soler. Linear-Scaling ab-initio Calculations for Large and Complex Systems. *Phys. Status Solidi B*, **215**, 809 (1999).

[11] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon. Simulated Quantum Computation of Molecular Energies. *Science*, **309**, 1704 (2005).

[12] A. Aspuru-Guzik and P. Walther. Photonic quantum simulators. *Nat. Phys.*, **8**, 285 (2012).

[13] R. Babbush, V. Denchev, N. Ding, S. Isakov, and H. Neven. Construction of non-convex polynomial loss functions for training a binary classifier with quantum annealing. *e-print arXiv: 1406.4203*, pages 1–15 (2014).

[14] R. Babbush, P. J. Love, and A. Aspuru-Guzik. Adiabatic Quantum Simulation of Quantum Chemistry. *Sci. Rep.*, **4**, 1 (2014).

[15] R. Babbush, J. McClean, D. Wecker, A. Aspuru-Guzik, and N. Wiebe. Chemical Basis of Trotter-Suzuki Errors in Chemistry Simulation. *Phys. Rev. A*, **91**, 022311 (2015).

[16] R. Babbush, B. O'Gorman, and A. A. Aspuru-Guzik. Resource efficient gadgets for compiling adiabatic quantum optimization problems. *Ann. Phys.*, **525**, 877 (2013).

[17] R. Babbush, J. Parkhill, and A. Aspuru-Guzik. Force-Field Functor Theory: Classical Force-Fields which Reproduce Equilibrium Quantum Distributions. *Front. Chem.*, **1**, 1 (2013).

[18] R. Babbush, A. Perdomo, B. O'Gorman, W. Macready, and A. Aspuru-Guzik. Construction of Energy Functions for Lattice Heteropolymer Models: Efficient Encodings for Constraint Satisfaction Programming and Quantum Annealing. In *Adv. Chem. Phys.*, volume 155, pages 201–243. 2014.

[19] R. Backofen. Using Constraint Programming for lattice Protein Folding. *Energy*, **3**, 389 (1998).

[20] R. Backofen and S. Will. A Constraint-Based Approach to Fast and Exact Structure Prediction in Three-Dimensional Protein Models. *Constraints*, **11**, 5 (2006).

[21] D. Baker. A surprising simplicity to protein folding. *Nature*, **405**, 39 (2000).

[22] G. Balasubramanian, P. Neumann, D. Twitchen, M. Markham, R. Kolesov, N. Mizuochi, J. Isoya, J. Achard, J. Beck, J. Tissler, V. Jacques, P. R. Hemmer, F. Jelezko, and J. Wrachtrup. Ultralong spin coherence time in isotopically engineered diamond. *Nat. Mater.*, **8**, 383 (2009).

[23] F. Barahona. On the computational complexity of Ising spin glass models. *J. Phys. A. Math. Gen.*, **15**, 3241 (1982).

[24] A. Barenco, C. Bennett, R. Cleve, D. DiVincenzo, N. Margolus, P. Shor, T. Sleator, J. Smolin, and H. Weinfurter. Elementary gates for quantum computation. *Phys. Rev. A*, **52**, 3457 (1995).

[25] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, Y. Campbell Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. Martinis. Superconducting quantum circuits at the surface code threshold for fault tolerance. *Nature*, **508**, 500 (2014).

[26] C. L. Benavides-Riveros, J. M. Garcia-Bondia, and M. Springborg. Why doubly excited determinants govern configuration interaction calculations of electron correlations. *e-print arXiv: 1409.6435* (2014).

[27] B. Berger and T. Leighton. Protein folding in the hydrophobic-hydrophilic (HP) model is NP-complete. *J. Comput. Biol. a J. Comput. Mol. cell Biol.*, **5**, 27 (1998).

[28] A. J. Berkley, M. W. Johnson, P. Bunyk, R. Harris, J. Johansson, T. Lanting, E. Ladizinsky, E. Tolkacheva, M. H. S. Amin, and G. Rose. A scalable readout system for a superconducting adiabatic quantum optimization system. *Supercond. Sci. Technol.*, **23**, 105014 (2010).

[29] A. J. Berkley, A. J. Przybysz, T. Lanting, R. Harris, N. Dickson, F. Altomare, M. H. Amin, P. Bunyk, C. Enderud, E. Hoskinson, M. W. Johnson, E. Ladizinsky, R. Neufeld, C. Rich, A. Y. Smirnov, E. Tolkacheva, S. Uchaikin, and A. B. Wilson. Tunneling spectroscopy using a probe qubit. *Phys. Rev. B*, **87**, 020502 (2013).

[30] A. Bermudez, F. Jelezko, M. B. Plenio, and A. Retzker. Electron-Mediated Nuclear-Spin Interactions between Distant Nitrogen-Vacancy Centers. *Phys. Rev. Lett.*, **107**, 150503 (2011).

[31] H. Bernien, B. Hensen, W. Pfaff, G. Koolstra, M. S. Blok, L. Robledo, T. H. Taminiau, M. Markham, D. J. Twitchen, L. Childress, and R. Hanson. Heralded entanglement between solid-state qubits separated by three metres. *Nature*, **497**, 86 (2013).

[32] E. Bernstein and U. Vazirani. Quantum complexity theory. *SIAM J. Comput.*, **26**, 1411 (1997).

[33] D. W. Berry, G. Ahokas, R. Cleve, and B. C. Sanders. Efficient Quantum Algorithms for Simulating Sparse Hamiltonians. *Commun. Math. Phys.*, **270**, 359 (2007).

[34] D. W. Berry and A. M. Childs. Black-box hamiltonian simulation and unitary implementation. *Quantum Inf. Comput.*, **12**, 29 (2012).

[35] D. W. Berry, A. M. Childs, R. Cleve, R. Kothari, and R. D. Somma. Exponential improvement in precision for simulating sparse Hamiltonians. In *STOC '14 Proc. 46th Annu. ACM Symp. Theory Comput.*, pages 283–292, 2014.

[36] J. D. Biamonte. Non-perturbative k-body to two-body commuting conversion Hamiltonians and embedding problem instances into Ising spins. *Phys. Rev. A*, **77**, 52331 (2008).

[37] J. D. Biamonte, V. Bergholm, J. Fitzsimons, and A. Aspuru-Guzik. Adiabatic quantum simulators. *AIP Adv.*, 1 (2011).

[38] J. D. Biamonte and P. J. Love. Realizable Hamiltonians for Universal Adiabatic Quantum Computers. *Phys. Rev. A*, **78**, 1 (2007).

[39] Z. Bian, F. Chudak, W. G. Macready, L. Clark, and F. Gaitan. Experimental Determination of Ramsey Numbers. *Phys. Rev. Lett.*, **111**, 130505 (2013).

[40] C. M. Bishop. Pattern Recognition and Machine Learning (Information Science and Statistics). 2006, 2006.

[41] A. Bocharov, M. Roetteler, and K. M. Svore. Efficient synthesis of probabilistic quantum circuits with fallback. *e-print arXiv: 1409.3552* (2014).

[42] B. M. Boghosian and W. Taylor. Simulating quantum mechanics on a quantum computer. *Phys. D-Nonlinear Phenom.*, **120**, 30 (1998).

[43] S. Boixo, T. Albash, F. M. Spedalieri, N. Chancellor, and D. A. Lidar. Experimental signature of programmable quantum annealing. *Nat. Commun.*, **4**, 2067 (2013).

[44] S. Boixo, T. F. Ronnow, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer. Evidence for quantum annealing with more than one hundred qubits. *Nat. Phys.*, **10**, 218 (2014).

[45] S. Boixo, V. N. Smelyanskiy, A. Shabani, S. V. Isakov, M. Dykman, V. S. Denchev, M. Amin, A. Smirnov, M. Mohseni, and H. Neven. Computational Role of Multiqubit Tunneling in a Quantum Annealer. *e-print arXiv: 1502.05754*, pages 1–7 (2015).

[46] S. Boixo and R. D. Somma. Necessary Condition for the Quantum Adiabatic Approximation. *Phys. Rev. A*, **81**, 5 (2009).

[47] B. Bollobás. *Modern Graph Theory*. Springer, New York, 1998.

[48] M. Born and V. Fock. Beweis des Adiabatensatzes. *Zeitschrift für Phys. A*, **51**, 165 (1928).

[49] E. Boros and A. Gruber. On Quadratization of Pseudo-Boolean Functions. *e-print arXiv:1404.6538*, pages 1–11 (2014).

[50] E. Boros and P. L. Hammer. Pseudo-boolean optimization. *Discret. Appl. Math.*, **123**, 155 (2002).

[51] D. Boughaci and H. Drias. Solving weighted Max-Sat optimization problems using a Taboo Scatter Search metaheuristic. *Proc. 2004 ACM Symp. Appl. Comput. SAC 04*, page 35 (2004).

[52] D. R. Bowler and T. Miyazaki. O(N) methods in electronic structure calculations. *Rep. Prog. Phys.*, **75**, 36503 (2012).

[53] S. Bravyi, D. DiVincenzo, and D. Loss. Schrieffer-Wolff transformation for quantum many-body systems. *Ann. Phys.*, 326 (2011).

[54] S. Bravyi, D. P. DiVincenzo, D. Loss, and B. M. Terhal. Quantum simulation of many-body Hamiltonians using perturbation theory with bounded-strength interactions. *Phys. Rev. Lett.*, **101**, 070503 (2008).

[55] S. Bravyi, D. P. DiVincenzo, R. I. Oliveira, and B. M. Terhal. The Complexity of Stoquastic Local Hamiltonian Problems. *Quantum Inf. Comput.*, **8**, 361 (2008).

[56] S. Bravyi and A. Kitaev. Fermionic quantum computation. *Ann. Phys. (N. Y).*, **298**, 210 (2002).

[57] J. Brooke, D. Bitko, T. Rosenbaum, and G. Aeppli. Quantum annealing of a disordered magnet. *Sci. (New York, NY)*, **284**, 779 (1999).

[58] K. L. Brown, W. J. Munro, and V. M. Kendon. Using Quantum Computers for Quantum Simulation. *Entropy*, **12**, 2268 (2010).

[59] I. Buluta and F. Nori. Quantum Simulators. *Science*, **326**, 108 (2009).

[60] P. I. Bunyk, E. M. Hoskinson, M. W. Johnson, E. Tolkacheva, F. Altomare, A. J. Berkley, R. Harris, J. P. Hilton, T. Lanting, A. J. Przybysz, and J. Whittaker. Architectural Considerations in the Design of a Superconducting Quantum Annealing Processor. *IEEE Trans. Appl. Supercond.*, **24**, 1 (2014).

[61] J. Cai, W. G. Macready, and A. Roy. A practical heuristic for finding graph minors. *e-print arXiv:1406.2741*, pages 1–16 (2014).

[62] C. J. Camacho. Entropic Barriers, Frustration and Order: Basic Ingredients in Protein Folding. *Phys. Rev. Lett.*, **77**, 4 (1995).

[63] Y. Cao, R. Babbush, J. Biamonte, and S. Kais. Hamiltonian Gadgets with Reduced Resource Requirements. *Phys. Rev. A*, **91**, 012315 (2015).

[64] R. Car and M. Parrinello. Unified Approach for Molecular Dynamics and Density-Functional Theory. *Phys. Rev. Lett.*, **55**, 2471 (1985).

[65] D. M. Ceperley. The statistical error of green's function Monte Carlo. *J. Stat. Phys.*, **43**, 815 (1986).

[66] K. Chandrasekaran, R. Karp, E. Moreno-Centeno, and S. Vempala. Algorithms for Implicit Hitting Set Problems. *SODA*, pages 614–629 (2011).

[67] D. Cheung, P. Hoyer, and N. Wiebe. Improved error bounds for the adiabatic approximation. *J. Phys. A Math. Theor.*, **44**, 415302 (2011).

[68] D. M. Chickering. Learning Bayesian networks is NP-complete. *Learn. from data*, pages 121–130 (1996).

[69] A. M. Childs, R. Cleve, E. Deotto, E. Farhi, S. Gutmann, and D. A. Spielman. Exponential algorithmic speedup by a quantum walk. *Proc. thirty-fifth Annu. ACM Symp. Theory Comput.*, **35**, 59 (2003).

[70] A. M. Childs, D. Gosset, and Z. Webb. The Bose-Hubbard model is QMA-complete. *e-print arXiv:1311.3297* (2013).

[71] A. M. Childs and N. Wiebe. Product formulas for exponentials of commutators. *J. Math. Phys.*, **54**, 062202 (2013).

[72] A. Choi, T. Standley, and A. Darwiche. Approximating Weighted Max-SAT Problems by Compensating for Relaxations. *CP*, pages 211–225 (2009).

[73] V. Choi. Minor-embedding in adiabatic quantum computation: I. The parameter setting problem. *Quantum Inf. Process.*, **7**, 193 (2008).

[74] V. Choi. Adiabatic Quantum Algorithms for the NP-Complete Maximum-Weight Independent Set, Exact Cover and 3SAT Problems. *e-print arXiv:1004.2226* (2010).

[75] V. Choi. Minor-embedding in adiabatic quantum computation: II. Minor-universal graph design. *Quantum Inf. Process.*, **10**, 343 (2011).

[76] V. Chvatal. A Greedy Heuristic for the Set-Covering Problem. *Math. Oper. Res.*, **4**, 233 (1979).

[77] N. Cody Jones, J. D. Whitfield, P. L. McMahon, M.-H. Yung, R. V. Meter, A. Aspuru-Guzik, and Y. Yamamoto. Faster quantum chemistry simulation on fault-tolerant quantum computers. *New J. Phys.*, **14**, 115023 (2012).

[78] I. Cohen and A. Retzker. Proposal for Verification of the Haldane Phase Using Trapped Ions. *Phys. Rev. Lett.*, **112**, 040503 (2014).

[79] S. A. Cook. The complexity of theorem-proving procedures. *Proc. third Annu. ACM Symp. Theory Comput. STOC '71*, **50**, 151 (1971).

[80] T. D. Crawford, C. D. Sherrill, E. F. Valeev, J. T. Fermann, R. A. King, M. L. Leininger, S. T. Brown, C. L. Janssen, E. T. Seidl, J. P. Kenny, and W. D. Allen. PSI3: An open-source Ab Initio electronic structure package. *J. Comput. Chem.*, **28**, 1610 (2007).

[81] P. Crescenzi, D. Goldman, C. Papadimitriou, A. Piccolboni, and M. Yannakakis. On the complexity of protein folding. *J. Comput. Biol. a J. Comput. Mol. cell Biol.*, **5**, 423 (1998).

[82] T. Cubitt and A. Montanaro. Complexity classification of local Hamiltonian problems. *Found. Comput. Sci. (FOCS), 2014 IEEE 55th Annu. Symp.*, pages 120 – 129 (2014).

[83] J. Cussens. Bayesian network learning by compiling to weighted MAX-SAT. In *UAI*, pages 105–112, 2008.

[84] A. Dal Palù, A. Dovier, and F. Fogolari. Constraint Logic Programming approach to protein structure prediction. *BMC Bioinformatics*, **5**, 186 (2004).

[85] V. S. Denchev. *Binary Classification with Adiabatic Quantum Optimization.* PhD thesis, Purdue University, 2013.

[86] V. S. Denchev, N. Ding, S. V. N. Vishwanathan, and H. Neven. Robust Classification with Adiabatic Quantum Optimization. *e-print arXiv: 1205.1148* (2012).

[87] 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. *Nat. Commun.*, **4**, 1903 (2013).

[88] K. A. Dill. Folding Proteins - Finding A Needle in A Haystack. *Curr. Opin. Struct. Biol.*, **3**, 99 (1993).

[89] K. A. Dill. Simple lattice models of protein folding. *Polym. Prepr. Am. Chem. Soc. Div. Polym. Chem.*, **36**, 635 (1995).

[90] K. A. Dill, S. B. Ozkan, M. S. Shell, and T. R. Weikl. The protein folding problem. *Annu. Rev. Biophys.*, **37**, 289 (2008).

[91] N. Ding, Y. Fang, R. Babbush, C. Chen, R. D. Skeel, and H. Neven. Bayesian Sampling Using Stochastic Gradient Thermostats. In *Adv. Neural Inf. Process. Syst.*, pages 3203–3211, 2014.

[92] P. Dirac. Quantum Mechanics of Many-Electron Systems. *Proc. R. Soc. London Ser. A, Contain. Pap. a Math. Phys. Character*, **123**, 714 (1929).

[93] A. Djebbari and J. Quackenbush. Seeded Bayesian Networks: Constructing genetic networks from microarray data. *BMC Syst. Biol.*, **2**, 1 (2008).

[94] M. W. Doherty, V. V. Struzhkin, D. A. Simpson, L. P. McGuinness, Y. Meng, A. Stacey, T. J. Karle, R. J. Hemley, N. B. Manson, L. Hollenberg Lloyd C., and S. Prawer. Electronic Properties and Metrology Applications of the Diamond NV-Center under Pressure. *Phys. Rev. Lett.*, **112**, 47601 (2014).

[95] F. Dolde, V. Bergholm, Y. Wang, I. Jakobi, B. Naydenov, S. Pezzagna, J. Meijer, F. Jelezko, P. Neumann, T. Schulte-Herbrüggen, J. Biamonte, and J. Wrachtrup. High-fidelity spin entanglement using optimal control. *Nat. Commun.*, 5 (2014).

[96] F. Dolde, I. Jakobi, B. Naydenov, N. Zhao, S. Pezzagna, C. Trautmann, J. Meijer, P. Neumann, F. Jelezko, and J. Wrachtrup. Room-temperature entanglement between single defect spins in diamond. *Nat. Phys.*, **8**, 1 (2013).

[97] J. Du, N. Xu, X. Peng, P. Wang, S. Wu, and D. Lu. NMR implementation of a molecular hydrogen quantum simulation with adiabatic state preparation. *Phys. Rev. Lett.*, **104**, 030502 (2010).

[98] Q.-H. Duan and P.-X. Chen. Realization of Universal Adiabatic Quantum Computation with Fewer Physical Resources. *Phys. Rev. A*, **84**, 4 (2011).

[99] M. V. G. Dutt, L. Childress, L. Jiang, E. Togan, J. Maze, F. Jelezko, a. S. Zibrov, P. R. Hemmer, and M. D. Lukin. Quantum register based on individual electronic and nuclear spin qubits in diamond. *Science*, **316**, 1312 (2007).

[100] N. Eén and N. Sörensson. Translating Pseudo-Boolean Constraints into SAT. *J. Satisf. Boolean Model. Comput.*, **2**, 1 (2006).

[101] A. E. Engel, N. Doss, G. J. Harris, and J. Tennyson. Calculated spectra for HeH+ and its effect on the opacity of cool metal-poor stars. *Mon. Not. R. Astron. Soc.*, **357**, 471 (2005).

[102] S. Even, A. Itai, and A. Shamir. On the Complexity of Timetable and Multicommodity Flow Problems. *SIAM J. Comput.*, **5**, 691 (1976).

[103] 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*, **292**, 472 (2001).

[104] E. Farhi, J. Goldstone, S. Gutmann, and M. Sipser. Quantum Computation by Adiabatic Evolution. *e-print arXiv: 0001106* (2000).

[105] E. Farhi, D. Gosset, I. Hen, A. Sandvik, A. Young, P. Shor, and F. Zamponi. Performance of the quantum adiabatic algorithm on random instances of two optimization problems on regular hypergraphs. *Phys. Rev. A*, 86 (2012).

[106] V. Feldman, V. Guruswami, P. Raghavendra, and Y. Wu. Agnostic learning of monomials by halfspaces is hard. *SIAM J. Comput.*, **41**, 1558 (2012).

[107] R. P. Feynman. Simulating physics with computers. *Int. J. Theor. Phys.*, **21**, 467 (1982).

[108] A. Finnila, M. Gomez, C. Sebenik, C. Stenson, and J. Doll. Quantum annealing: A new method for minimizing multidimensional functions. *Chem. Phys. Lett.*, **219**, 343 (1994).

[109] A. Friedenauer, H. Schmitz, J. T. Glueckert, D. Porras, and T. Schaetz. Simulating a quantum magnet with trapped ions. *Nat. Phys.*, **4**, 757 (2008).

[110] N. Friedman, M. Goldszmidt, and A. Wyner. Data Analysis with Bayesian Networks: A Bootstrap Approach. In *Proc. Fifteenth Conf. Uncertain. Artif. Intell.*, UAI'99, pages 196–205, San Francisco, CA, USA, 1999. Morgan Kaufmann Publishers Inc.

[111] N. Friedman, M. Linial, I. Nachman, and D. Pe'er. Using Bayesian Networks to Analyze Expression Data. In *Proc. Fourth Annu. Int. Conf. Comput. Mol. Biol.*, RECOMB '00, pages 127–135, New York, NY, USA, 2000. ACM.

[112] F. Gaitan and L. Clark. Graph isomorphism and adiabatic quantum computing. *Phys. Rev. A*, **89**, 22342 (2013).

[113] A. Ganti, U. Onunkwo, and K. Young. A family of [[6k, 2k, 2]] codes for practical, scalable adiabatic quantum computation. *e-print arXiv:1309.1674* (2013).

[114] M. R. Garey and D. S. Johnson. *Computers and Intractability: A Guide to the Theory of NP-Completeness.* W. H. Freeman, 1979.

[115] S. Garnerone, P. Zanardi, and D. A. Lidar. Adiabatic Quantum Algorithm for Search Engine Ranking. *Phys. Rev. Lett.*, **108**, 230506 (2012).

[116] I. Georgescu, S. Ashhab, and F. Nori. Quantum simulation. *Rev. Mod. Phys.*, **86**, 153 (2014).

[117] I. Georgescu, S. E. Brown, and V. A. Mandelshtam. Mapping the phase diagram for neon to a quantum Lennard-Jones fluid using Gibbs ensemble simulations. *J. Chem. Phys.*, 138 (2013).

[118] E. Gibney. Physics: Quantum computer quest. *Nature*, **516**, 24 (2014).

[119] B. Giles and P. Selinger. Exact synthesis of multiqubit Clifford + T circuits. *Phys. Rev. A*, **87**, 32332 (2013).

[120] J. I. Gillen, W. S. Bakr, A. Peng, P. Unterwaditzer, S. Fölling, and M. Greiner. Two-dimensional quantum gas in a hybrid surface trap. *Phys. Rev. A*, **80**, 21602 (2009).

[121] S. Goedecker. Linear scaling electronic structure methods. *Rev. Mod. Phys.*, **71**, 1085 (1999).

[122] M. Gordon and M. Schmidt. Advances in electronic structure theory: GAMESS a decade later. In C. Dykstra, G. Frenking, K. Kim, and G. Scuseria, editors, *Theory Appl. Comput. Chem. first forty years*, pages 1167–1189. Elsevier, Amersterdam, 2005.

[123] D. Gosset and D. Nagaj. Quantum 3-SAT is QMA1-complete. *e-print arXiv:1302.0290*, pages 1–44 (2013).

[124] L. K. Grover. A fast quantum mechanical algorithm for database search. In *Proc. twenty-eighth Annu. ACM Symp. Theory Comput.*, STOC '96, pages 212–219, New York, NY, USA, 1996. ACM.

[125] A. Gruber, A. Drabenstedt, C. Tietz, L. Fleury, J. Wrachtrup, and C. VonBorczyskowski. Scanning confocal optical microscopy and magnetic resonance on single defect centers. *Science*, **276**, 2012 (1997).

[126] M. Gruebele and P. G. Wolynes. Satisfying turns in folding transitions., 1998.

[127] J. Hachmann, W. Cardoen, and G. K.-L. Chan. Multireference correlation in long molecules with the quadratic scaling density matrix renormalization group. *J. Chem. Phys.*, **125**, (2006).

[128] J. P. Hague, S. Downes, C. MacCormick, and P. E. Kornilovitch. Cold Rydberg atoms for quantum simulation of exotic condensed matter interactions. *J. Supercond. Nov. Magn.* (2013).

[129] B. L. Hammond, P. J. Reynolds, and W. A. Lester. Valence quantum Monte Carlo with ab initio effective core potentials. *J. Chem. Phys.*, 87 (1987).

[130] P. Hansen and B. Jaumard. Approximation Algorithms for the Maximum Satisfiability Problem. *Computing*, **7**, 279 (1990).

[131] P. Hansen, B. Jaumard, and M. P. De Aragao. Mixed-integer column generation algorithms and the probabilistic maximum satisfiability problem. *Eur. J. Oper. Res.*, **108**, 671 (1998).

[132] R. Harris, A. J. Berkley, M. W. Johnson, P. Bunyk, S. Govorkov, M. C. Thom, S. Uchaikin, A. Wilson, J. Chung, and E. Holtham. Sign-and magnitude-tunable coupler for superconducting flux qubits. *Phys. Rev. Lett.*, **98**, 177001 (2007).

[133] R. Harris, F. Brito, A. J. Berkley, J. Johansson, M. W. Johnson, T. Lanting, P. Bunyk, E. Ladizinsky, B. Bumble, and A. Fung. Synchronization of multiple coupled rf-SQUID flux qubits. *New J. Phys.*, **11**, 123022 (2009).

[134] R. Harris, M. W. Johnson, S. Han, A. J. Berkley, J. Johansson, P. Bunyk, E. Ladizinsky, S. Govorkov, M. C. Thom, S. Uchaikin, B. Bumble, A. Fung, A. Kaul, A. Kleinsasser, M. H. S. Amin, and D. V. Averin. Probing Noise in Flux Qubits via Macroscopic Resonant Tunneling. *Phys. Rev. Lett.*, **101**, 117003 (2008).

[135] 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. *Phys. Rev. B*, **82**, 16 (2010).

[136] R. Harris, T. Lanting, A. J. Berkley, J. Johansson, M. W. Johnson, P. Bunyk, E. Ladizinsky, N. Ladizinsky, T. Oh, and S. Han. Compound Josephson-junction coupler for flux qubits with minimal crosstalk. *Phys. Rev. B*, **80**, 52506 (2009).

[137] R. J. Harrison, G. I. Fann, T. Yanai, Z. Gan, and G. Beylkin. Multiresolution quantum chemistry: Basic theory and initial applications. *J. Chem. Phys.*, 121 (2004).

[138] A. W. Harrow. Why now is the right time to study quantum computing. *e-print arXiv: 1501.00011*, pages 1–6 (2015).

[139] W. E. Hart and S. Istrail. Robust proofs of NP-hardness for protein folding: general lattices and energy potentials. *J. Comput. Biol. a J. Comput. Mol. cell Biol.*, **4**, 1 (1997).

[140] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer. Improving Quantum Algorithms for Quantum Chemistry. *Quantum Inf. Comput.*, **15**, 1 (2015).

[141] P. Hauke, D. Marcos, M. Dalmonte, and P. Zoller. Quantum simulation of a lattice Schwinger model in a chain of trapped ions. *Phys. Rev. X*, **3**, 18 (2013).

[142] P. Hayden, D. W. Leung, and A. Winter. Aspects of generic entanglement. *Commun. Math. Phys.*, **265**, 95 (2006).

[143] D. Heckerman, D. Geiger, and D. M. Chickering. Learning Bayesian Networks: The Combination of Knowledge and Statistical Data. *Mach. Learn.*, **20**, 197 (1995).

[144] W. J. Hehre, R. F. Stewart, and J. A. Pople. Self-Consistent Molecular Orbital Methods I: Use of Gaussian Expansions of Slater-Type Atomic Orbitals. *J. Chem. Phys.*, **51**, 2657 (1969).

[145] T. Helgaker, P. Jorgensen, and J. Olsen. *Molecular Electronic Structure Theory.* Wiley, 2002.

[146] R. Hemmecke, M. Köppe, J. Lee, and R. Weismantel. Nonlinear Integer Programming. *50 Years Integer Program.*, page 57 (2009).

[147] I. Hen and A. Young. Solving the Graph Isomorphism Problem with a Quantum Annealer. *Phys. Rev. A*, 86 (2012).

[148] T. Hogg. Adiabatic quantum computing for random satisfiability problems. *Phys. Rev. A*, **67**, 22314 (2003).

[149] IBM. IBM ILOG CPLEX V12.1: User's Manual for CPLEX, 2009.

[150] S. V. Isakov, I. N. Zintchenko, T. F. Ronnow, and M. Troyer. Optimized simulated annealing code for Ising spin glasses. *e-print arXiv: 1401.1084*, pages 1–10 (2014).

[151] F. Jelezko, T. Gaebel, I. Popa, M. Domhan, A. Gruber, and J. Wrachtrup. Observation of coherent oscillation of a single nuclear spin and realization of a two-qubit conditional quantum gate. *Phys. Rev. Lett.*, **93**, (2004).

[152] M. Johanning, A. F. Varón, and C. Wunderlich. Quantum simulations with cold trapped ions. *J. Phys. B At.*, **42**, 4009 (2009).

[153] J. Johansson, M. H. S. Amin, A. J. Berkley, P. Bunyk, V. Choi, R. Harris, M. W. Johnson, T. M. Lanting, S. Lloyd, and G. Rose. Landau-Zener transitions in a superconducting flux qubit. *Phys. Rev. B*, **80**, 12507 (2009).

[154] 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*, **473**, 194 (2011).

[155] C. Jones. Low-overhead constructions for the fault-tolerant Toffoli gate. *Phys. Rev. A*, **87**, 022328 (2013).

[156] P. Jordan and E. Wigner. über das paulische äquivalenzverbot. *Zeitschrift für Phys.*, **47**, 631 (1928).

[157] S. P. Jordan and E. Farhi. Perturbative Gadgets at Arbitrary Orders. *Phys. Rev. A*, **77**, 062329 (2008).

[158] S. P. Jordan, D. Gosset, and P. J. Love. Quantum Merlin Arthur complete problems for stoquastic Hamiltonians and Markov matrices. *Phys. Rev. A*, **81**, 1 (2010).

[159] T. Kadowaki and H. Nishimori. Quantum Annealing in the Transverse Ising Model. *Phys. Rev. E*, **58**, 5355 (1998).

[160] F. Kahl and P. Strandmark. Generalized roof duality for pseudo-boolean optimization, 2011.

[161] W. M. Kaminsky and S. Lloyd. Scalable Architecture for Adiabatic Quantum Computing of Np-Hard Problems. In A. J. Leggett, B. Ruggiero, and P. Silvestrini, editors, *Quantum Comput. Quantum Bits Mesoscopic Syst.*, pages 229–236. Springer US, 2004.

[162] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik. Polynomial-time quantum algorithm for the simulation of chemical dynamics. *Proc. Natl. Acad. Sci.*, **105**, 18681 (2008).

[163] I. Kassal, J. Whitfield, A. Perdomo-Ortiz, M.-H. Yung, and A. Aspuru-Guzik. Simulating Chemistry Using Quantum Computers. *Ann. Rev. Phys. Chem.*, **62**, 185 (2010).

[164] H. G. Katzgraber, F. Hamze, and R. S. Andrist. Glassy Chimeras Could Be Blind to Quantum Speedup: Designing Better Benchmarks for Quantum Annealing Machines. *Phys. Rev. X*, **4**, 021008 (2014).

[165] J. Kempe, A. Kitaev, and O. Regev. The Complexity of the Local Hamiltonian Problem. *SIAM J. Comput.*, **35**, 1070 (2006).

[166] N. Khaneja, T. Reiss, C. Kehlet, T. Schulte-Herbruggen, and S. J. Glaser. Optimal control of coupled spin dynamics: design of NMR pulse sequences by gradient ascent algorithms. *J. Magn. Reson.*, **172**, 296 (2005).

[167] S. Kirkpatrick, M. P. Vecchi, and Others. Optimization by simmulated annealing. *Science*, **220**, 671 (1983).

[168] A. Y. Kitaev. Quantum measurements and the Abelian Stabilizer Problem. *e-print arXiv: 9511026*, pages 1–22 (1995).

[169] A. Y. Kitaev, A. H. Shen, and M. N. Vyalyi. *Classical and Quantum Computation*, volume 47. Amer Mathematical Society, 2002.

[170] V. Kliuchnikov, D. Maslov, and M. Mosca. Fast and efficient exact synthesis of single-qubit unitaries generated by Clifford and T gates. *Quantum Inf. Comput.*, **13**, 607 (2013).

[171] D. Koller and N. Friedman. *Probabilistic Graphical Models: Principles and Techniques - Adaptive Computation and Machine Learning*. The MIT Press, 2009.

[172] L. Krippahl and P. Barahona. Applying Constraint Programming to Protein Structure Determination. *Princ. Pract. Constraint Program. CP99*, **1713**, 289 (1999).

[173] T. Lanting, A. J. Berkley, B. Bumble, P. Bunyk, A. Fung, J. Johansson, A. Kaul, A. Kleinsasser, E. Ladizinsky, and F. Maibaum. Geometrical dependence of the low-frequency noise in superconducting flux qubits. *Phys. Rev. B*, **79**, 60509 (2009).

[174] T. Lanting, A. A. Przybysz, A. Y. Smirnov, F. F. Spedalieri, M. M. Amin, A. A. Berkley, R. Harris, F. Altomare, S. Boixo, P. Bunyk, N. Dickson, C. Enderud, J. J. Hilton, E. Hoskinson, M. M. Johnson, E. Ladizinsky, N. Ladizinsky, R. Neufeld, T. Oh, I. Perminov, C. Rich, M. M. Thom, E. Tolkacheva, S. Uchaikin, A. A. Wilson, and G. Rose. Entanglement in a Quantum Annealing Processor. *Phys. Rev. X*, **4**, 1 (2014).

[175] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and a. G. White. Towards quantum chemistry on a quantum computer. *Nat. Chem.*, **2**, 106 (2010).

[176] J. Larrosa, F. Heras, and S. De Givry. A Logical Approach to Efficient Max-SAT solving. *Artif. Intell.*, **172**, 204 (2006).

[177] K. F. Lau and K. A. Dill. A lattice statistical mechanics model of the conformational and sequence spaces of proteins. *Macromolecules*, **22**, 3986 (1989).

[178] S. Laue. Geometric Set Cover and Hitting Sets for Polytopes in R^3. *Science*, **2008**, 479 (2008).

[179] D. Leibfried, B. Demarco, V. Meyer, M. Rowe, A. Ben-Kish, J. Britton, W. M. Itano, B. Jelenković, C. Langer, T. Rosenband, and D. J. Wineland. Trapped-Ion Quantum Simulator: Experimental Application to Nonlinear Interferometers. *Phys. Rev. Lett.*, **89**, 247901 (2002).

[180] H. Li, Y. Liang, and Q. Xu. Support vector machines and its applications in chemistry. *Chemom. Intell. Lab. Syst.*, **95**, 188 (2009).

[181] Z. Li, M.-H. Yung, H. Chen, D. Lu, J. D. Whitfield, X. Peng, A. Aspuru-Guzik, and J. Du. Solving Quantum Ground-State Problems with Nuclear Magnetic Resonance. *Sci. Rep.*, **1**, 1 (2011).

[182] D. A. Lidar and O. Biham. Simulating Ising spin glasses on a quantum computer. *Phys. Rev. E (Statistical Phys.*, **56**, 3661 (1997).

[183] S. Lloyd. Universal Quantum Simulators. *Science*, **273**, 1073 (1996).

[184] P. M. Long and R. A. Servedio. Random classification noise defeats all convex potential boosters. *Mach. Learn.*, **78**, 287 (2010).

[185] P. Löwdin. On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals. *J. Chem. Phys.*, **18**, 365 (1950).

[186] D. Lu, B. Xu, N. Xu, Z. Li, H. Chen, X. Peng, R. Xu, and J. Du. Quantum chemistry simulation on quantum computers: theories and experiments. *Phys. Chem. Chem. Phys.*, **14**, 9411 (2012).

[187] D. Lu, N. Xu, R. Xu, H. Chen, J. Gong, X. Peng, and J. Du. Simulation of chemical isomerization reaction dynamics on a NMR quantum simulator. *Phys. Rev. Lett.*, **107**, 020501 (2011).

[188] X.-S. Ma, B. Dakić, S. Kropatsche, W. Naylor, Y.-h. Chan, Z.-x. Gong, L.-m. Duan, A. Zeilinger, and P. Walther. Photonic quantum simulation of ground state configurations of Heisenberg square and checkerboard lattice spin systems. *e-print arXiv:1205.2801* (2012).

[189] Y. Manin. Vychislimoe i nevychislimoe [Computable and Noncomputable]. *Sov. Radio.*, pages 13–15 (1980).

[190] P. Manolios and V. Papavasileiou. Pseudo-Boolean Solving by incremental translation to SAT, 2011.

[191] N. Manwani and P. S. Sastry. Noise tolerance under risk minimization. *Cybern. IEEE Trans.*, **43**, 1146 (2013).

[192] J. Marques-Silva and K. A. Sakallah. Theory and Applications of Satisfiability Testing - SAT 2007, 10th International Conference, Lisbon, Portugal, May 28-31, 2007, Proceedings. In J. a. Marques-Silva and K. A. Sakallah, editors, *SAT*, volume 4501 of *Lecture Notes in Computer Science*, page 384. Springer, 2007.

[193] P. C. Maurer, G. Kucsko, C. Latta, L. Jiang, N. Y. Yao, S. D. Bennett, F. Pastawski, D. Hunger, N. Chisholm, M. Markham, D. J. Twitchen, J. I. Cirac, and M. D. Lukin. Room-Temperature Quantum Bit Memory Exceeding One Second. *Science*, **336**, 1283 (2012).

[194] I. Mayer. On Löwdin's method of symmetric orthogonalization. *Int. J. Quant. Chem.*, **90**, 63 (2002).

[195] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik. Exploiting locality in quantum computation for quantum chemistry. *J. Phys. Chem. Lett.*, **5**, 4368 (2014).

[196] D. Mease and A. Wyner. Evidence contrary to the statistical view of boosting. *J. Mach. Learn. Res.*, **9**, 131 (2008).

[197] D. A. Meyer. From quantum cellular automata to quantum lattice gases. *J. Stat. Phys.*, **85**, 551 (1996).

[198] L. Mirny and E. Shakhnovich. Protein folding theory: from lattice to all-atom models. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 361 (2001).

[199] S. Miyazawa and R. L. Jernigan. Residue-residue potentials with a favorable contact pair term and an unfavorable high packing density term, for simulation and threading. *J. Mol. Biol.*, **256**, 623 (1996).

[200] A. Mizel, D. A. Lidar, and M. Mitchell. Simple Proof of Equivalence between Adiabatic Quantum Computation and the Circuit Model. *Phys. Rev. Lett.*, **99**, 70502 (2007).

[201] S. Morita and H. Nishimori. Convergence theorems for quantum annealing. *J. Phys. A. Math. Gen.*, **39**, 13903 (2006).

[202] S. Morita and H. Nishimori. Mathematical foundation of quantum annealing. *J. Math. Phys.*, **49**, 125210 (2008).

[203] D. Nagaj. Fast universal quantum computation with railroad-switch local Hamiltonians. *J. Math. Phys.*, **51**, 2201 (2010).

[204] D. Nagaj and S. Mozes. New construction for a QMA complete three-local Hamiltonian. *J. Math. Phys.*, **48**, 2104 (2007).

[205] T. Neuhaus, M. Peschina, K. Michielsen, and H. De Raedt. Classical and quantum annealing in the median of three-satisfiability. *Phys. Rev. A*, **83**, 12309 (2011).

[206] P. Neumann, J. Beck, M. Steiner, F. Rempp, H. Fedder, P. R. Hemmer, J. Wrachtrup, and F. Jelezko. Single-shot readout of a single nuclear spin. *Science*, **329**, 542 (2010).

[207] P. Neumann, N. Mizuochi, F. Rempp, P. Hemmer, H. Watanabe, S. Yamasaki, V. Jacques, T. Gaebel, F. Jelezko, and J. Wrachtrup. Multipartite entanglement among single spins in diamond. *Science*, **320**, 1326 (2008).

[208] H. Neven, V. Denchev, M. Drew-Brook, J. Zhang, W. G. Macready, G. Rose, and S. Aaronson. BQP and the Polynomial Hierarchy. In *Proc. Forty-second ACM Symp. Theory Comput.*, STOC '10, pages 141–150, New York, NY, USA, 2010. ACM.

[209] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready. Training a Binary Classifier with the Quantum Adiabatic Algorithm. *e-print arXiv: 0811.0416*, pages 1–11 (2008).

[210] H. Neven, V. S. Denchev, G. Rose, and W. G. Macready. Training a Large Scale Classifier with the Quantum Adiabatic Algorithm. *e-print arXiv:09120779*, page 14 (2009).

[211] H. Neven, G. Rose, and W. G. W. G. Macready. Image recognition with an adiabatic quantum computer I. Mapping to quadratic unconstrained binary optimization. *e-print arXiv:0804.4457* (2008).

[212] H. Nishimori. Relation of classical non-equilibrium dynamics and quantum annealing. *e-print arXiv: 1503.02127*, pages 1–5 (2015).

[213] H. Nymeyer, A. E. García, and J. N. Onuchic. Folding funnels and frustration in off-lattice minimalist protein landscapes. *Proc. Natl. Acad. Sci. U. S. A.*, **95**, 5921 (1998).

[214] M. T. Oakley, D. J. Wales, and R. L. Johnston. Energy Landscape and Global Optimization for a Frustrated Model Protein. *J. Phys. Chem. B*, **115**, 11525 (2011).

[215] B. O'Gorman, A. Perdomo-Ortiz, R. Babbush, A. Aspuru-Guzik, and V. Smelyanskiy. Bayesian Network Structure Learning Using Quantum Annealing. *Eur. Phys. J. Spec. Top.*, **225**, 163 (2015).

[216] R. Oliveira and B. M. Terhal. The complexity of quantum spin systems on a two-dimensional square lattice. *Quantum Inf. Comput.*, **8**, 900 (2008).

[217] G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme. Quantum algorithms for fermionic simulations. *Phys. Rev. A*, **64**, 022319 (2001).

[218] D. Pankratov and A. Borodin. On the Relative Merits of Simple Local Search Methods for the MAX-SAT Problem. In O. Strichman and S. Szeider, editors, *Theory Appl. Satisf. Test. SAT 2010*, volume 6175 of *Lecture Notes in Computer Science*, pages 223–236. Springer, 2010.

[219] S. Parker and M. Plenio. Efficient Factorization with a Single Pure Qubit and log N Mixed Qubits. *Phys. Rev. Lett.*, **85**, 3049 (2000).

[220] A. Perdomo, C. Truncik, I. Tubert-Brohman, G. Rose, and A. Aspuru-Guzik. Construction of model Hamiltonians for adiabatic quantum computation and its application to finding low-energy conformations of lattice protein models. *Phys. Rev. A*, **78**, 12320 (2008).

[221] A. Perdomo-Ortiz, N. Dickson, M. Drew-Brook, G. Rose, and A. Aspuru-Guzik. Finding low-energy conformations of lattice protein models by quantum annealing. *Sci. Rep.*, **2**, 1 (2012).

[222] A. Perdomo-Ortiz, J. Fluegemann, S. Narasimhan, V. Smelyanskiy, and R. Biswas. A quantum annealing approach for fault detection and diagnosis of graph-based systems. *Eur. Phys. J. Spec. Top.*, **224**, 131 (2014).

[223] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. OBrien. A variational eigenvalue solver on a photonic quantum processor. *Nat. Commun.*, **5**, 1 (2014).

[224] D. Poulin, M. B. Hastings, D. Wecker, N. Wiebe, A. C. Doherty, and M. Troyer. The Trotter Step Size Required for Accurate Quantum Simulation of Quantum Chemistry. *Quantum Inf. Comput.*, **15**, 361 (2015).

[225] K. L. Pudenz, T. Albash, and D. A. Lidar. Error-corrected quantum annealing with hundreds of qubits. *Nat. Commun.*, **5**, 3243 (2014).

[226] S. Raeisi, N. Wiebe, and B. C. Sanders. Quantum-circuit design for efficient simulations of many-body quantum dynamics. *New J. Phys.*, **14**, 3017 (2012).

[227] J. Rennie. On L2-norm regularization and the Gaussian prior. Technical report, MIT, 2003.

[228] P. Richerme, C. Senko, J. Smith, A. Lee, S. Korenblit, and C. Monroe. Experimental performance of a quantum simulator: Optimizing adiabatic evolution and identifying many-body ground states. *Phys. Rev. A*, **88**, 12334 (2013).

[229] E. G. Rieffel, D. Venturelli, B. O'Gorman, M. Do, E. Prystay, and V. Smelyanskiy. A case study in programming a quantum annealer for hard operational planning problems. *To be Submitt.* (2014).

[230] L. Robledo, L. Childress, H. Bernien, B. Hensen, P. F. a. Alkemade, and R. Hanson. High-fidelity projective read-out of a solid-state spin quantum register. *Nature*, **477**, 574 (2011).

[231] J. Roland and N. J. Cerf. Quantum search by local adiabatic evolution. *Phys. Rev. A*, **65**, 42308 (2002).

[232] J. Roland and N. J. Cerf. Adiabatic quantum search algorithm for structured problems. *Phys. Rev. A*, **68**, 62312 (2003).

[233] T. F. Ronnow, 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*, **345**, 420 (2014).

[234] N. J. Ross and P. Selinger. Optimal ancilla-free Clifford + T approximation of z-rotations. *e-print arXiv: 1403.2975* (2014).

[235] K. A. Sakallah. Pueblo : A Hybrid Pseudo-Boolean SAT Solver. *Electr. Eng.*, **2**, 155 (2006).

[236] A. Sali, E. Shakhnovich, and M. Karplus. How does a protein fold? *Nature*, **369**, 248 (1994).

[237] B. C. Sanders. Efficient Algorithms for Universal Quantum Simulation. *Lect. Notes Comput. Sci.*, **7948**, 1 (2013).

[238] G. E. Santoro, R. Martovnák, E. Tosatti, and R. Car. Theory of quantum annealing of an Ising spin glass. *Science*, **295**, 2427 (2002).

[239] G. E. Santoro and E. Tosatti. Optimization using quantum mechanics: quantum annealing through adiabatic evolution. *J. Phys. A. Math. Gen.*, **39**, R393 (2006).

[240] M. Schmidt, K. Baldridge, J. Boatz, S. Elbert, M. Gordon, J. Jensen, S. Koeski, N. Matsunaga, K. Nguyen, S. Su, T. Windus, M. Dupuis, and J. Montgomery. General Atomic and Molecular Electronic Structure System. *J. Comput. Chem.*, **14**, 1347 (1993).

[241] R. D. Schram, G. T. Barkema, and R. H. Bisseling. Exact enumeration of self-avoiding walks. *J. Stat. Mech. Theory Exp.*, **2011**, 6 (2011).

[242] N. Schuch and F. Verstraete. Computational complexity of interacting electrons and fundamental limitations of density functional theory. *Nat. Phys.*, **5**, 732 (2009).

[243] J. T. Seeley, M. J. Richard, and P. J. Love. The Bravyi-Kitaev transformation for quantum computation of electronic structure. *J. Chem. Phys.*, **137**, 224109 (2012).

[244] P. Selinger. Quantum circuits of T-depth one. *Phys. Rev. A*, **87**, 042302 (2013).

[245] E. I. Shakhnovich. Proteins with selected sequences fold into unique native conformation. *Phys. Rev. Lett.*, **72**, 3907 (1994).

[246] E. I. Shakhnovich. Modeling protein folding: the beauty and power of simplicity. *Fold. Des.*, **1**, R50 (1996).

[247] Y. Shao, L. F. Molnar, Y. Jung, J. Kussmann, C. Ochsenfeld, S. T. Brown, A. T. B. Gilbert, L. V. Slipchenko, S. V. Levchenko, D. P. O'Neill, and Others. Advances in methods and algorithms in a modern quantum chemistry program package. *Phys. Chem. Chem. Phys.*, **8**, 3172 (2006).

[248] J.-E. Shea, J. N. Onuchic, and C. L. Brooks. Energetic frustration and the nature of the transition state in protein folding. *J. Chem. Phys.*, **113**, 7663 (2000).

[249] H. M. Sheini and K. A. Sakallah. Pueblo: a modern pseudo-Boolean SAT solver, 2005.

[250] L. Shi and X. Cai. An Exact Fast Algorithm for Minimum Hitting Set, 2010.

[251] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. *Proc. 35th Annu. Symp. Found. Comput. Sci.*, pages 124–134 (1994).

[252] J. Simon, W. S. Bakr, R. Ma, M. E. Tai, P. M. Preiss, and M. Greiner. Quantum simulation of antiferromagnetic spin chains in an optical lattice. *Nature*, **472**, 307 (2011).

[253] V. N. Smelyanskiy, E. G. Rieffel, S. I. Knysh, C. P. Williams, M. W. Johnson, M. C. Thom, W. G. Macready, and K. L. Pudenz. A Near-Term Quantum Computing Approach for Hard Computational Problems in Space Exploration. *Electr. Eng.*, page 68 (2012).

[254] J. A. Smolin and G. Smith. Classical signatures of quantum annealing. *e-print arXiv:1305.4904*, pages 1–8 (2013).

[255] R. D. Somma, S. Boixo, H. Barnum, and E. Knill. Quantum simulations of classical annealing processes. *Phys. Rev. Lett.*, **101**, 130504 (2008).

[256] R. D. Somma, D. Nagaj, and M. Kieferová. Quantum Speedup by Quantum Annealing. *Phys. Rev. Lett.*, **109**, 50501 (2012).

[257] R. D. Somma, G. Ortiz, J. Gubernatis, E. Knill, and R. Laflamme. Simulating physical phenomena by quantum networks. *Phys. Rev. A*, **65**, 17 (2002).

[258] M. Soos, K. Nohl, and C. Castelluccia. Extending SAT Solvers to Cryptographic Problems. *SAT*, **5584**, 244 (2009).

[259] M. Suzuki. General Decomposition Theory of Ordered Exponentials. *Proc. Jpn. Acad. Ser. B Phys. Biol. Sci.*, **69**, 161 (1993).

[260] T. H. Taminiau, J. Cramer, T. van der Sar, V. V. Dobrovitski, and R. Hanson. Universal control and error correction in multi-qubit spin registers in diamond. *Nat. Nanotechnol.*, pages 1–7 (2014).

[261] J. M. Taylor, P. Cappellaro, L. Childress, L. Jiang, D. Budker, P. R. Hemmer, A. Yacoby, R. Walsworth, and M. D. Lukin. High-sensitivity diamond magnetometer with nanoscale resolution. *Nat. Phys.*, **4**, 29 (2008).

[262] L. Thogersen and J. Olsen. A coupled cluster and full configuration interaction study of CN and CN-. *Chem. Phys. Lett.*, **393**, 36 (2004).

[263] E. Togan, Y. Chu, A. S. Trifonov, L. Jiang, J. Maze, L. Childress, M. V. G. Dutt, A. S. Sorensen, P. R. Hemmer, A. S. Zibrov, and M. D. Lukin. Quantum entanglement between an optical photon and a solid-state spin qubit. *Nature*, **466**, 730 (2010).

[264] B. Toloui and P. J. Love. Quantum Algorithms for Quantum Chemistry based on the sparsity of the CI-matrix. *e-print arXiv:1312.2579*, pages 1–20 (2013).

[265] A. Trabesinger. Quantum simulation. *Nat. Phys.*, **8**, 263 (2012).

[266] H. F. Trotter. On the product of semi-groups of operators. *Proc. Am. Math. Soc.*, **10**, 545 (1959).

[267] R. R. Tucci. An introduction to quantum Bayesian networks for mixed states. *e-print arXiv: 1204.1550*, pages 1–20 (2012).

[268] R. R. Tucci. Quantum circuit for discovering from data the structure of classical Bayesian networks. *e-print arXiv: 1404.0055*, pages 1–20 (2014).

[269] A. D. Ullah and K. Steinhöfel. A hybrid approach to protein folding problem integrating constraint programming with local search. *BMC Bioinformatics*, **11 Suppl 1**, S39 (2010).

[270] J. H. van Vleck. Nonorthogonality and Ferromagnetism. *Phys. Rev.*, **49**, 232 (1936).

[271] L. Veis and J. Pittner. Quantum computing applied to calculations of molecular energies: CH2 benchmark. *J. Chem. Phys.*, **133**, 194106 (2010).

[272] L. Veis and J. Pittner. Adiabatic state preparation study of methylene. *J. Chem. Phys.*, **140**, 1 (2014).

[273] D. Venturelli, S. Mandrà, S. Knysh, B. O'Gorman, R. Biswas, and V. Smelyanskiy. Quantum Optimization of Fully-Connected Spin Glasses. *e-print arXiv: 1406.7553*, pages 1–14 (2014).

[274] F. Verstraete and J. I. Cirac. Mapping local Hamiltonians of fermions to local Hamiltonians of spins. *J. Stat. Mech. Theory Exp.*, page P09012 (2005).

[275] G. Waldherr, Y. Wang, S. Zaiser, M. Jamali, T. Schulte-Herbrüggen, H. Abe, T. Ohshima, J. Isoya, J. F. Du, P. Neumann, and J. Wrachtrup. Quantum error correction in a solid-state hybrid spin register. *Nature* (2014).

[276] H. Wang, S. Ashhab, and F. Nori. Efficient quantum algorithm for preparing molecular-system-like states on a quantum computer. *Phys. Rev. A - At. Mol. Opt. Phys.*, 79 (2009).

[277] H. Wang, S. Kais, A. Aspuru-Guzik, and M. R. Hoffmann. Quantum algorithm for obtaining the energy spectrum of molecular systems. *Phys. Chem. Chem. Phys.*, **10**, 5388 (2008).

[278] L. Wang, T. F. Ronnow, S. Boixo, S. V. Isakov, Z. Wang, D. Wecker, D. A. Lidar, J. M. Martinis, and M. Troyer. Comment on: "Classical signature of quantum annealing". *e-print arXiv:1305.5837*, pages 1–3 (2013).

[279] Y. Wang, F. Dolde, J. Biamonte, R. Babbush, V. Bergholm, S. Yang, I. Jakobi, P. Neumann, A. Aspuru-Guzik, J. D. Whitfield, and J. Wrachtrup. Quantum Simulation of Helium Hydride in a Solid-State Spin Register. *e-print arXiv: 1405.2696*, pages 1–9 (2014).

[280] N. J. Ward, I. Kassal, and A. Aspuru-Guzik. Preparation of many-body states for quantum simulation. *J. Chem. Phys.*, **130**, 194105 (2008).

[281] D. Wecker, B. Bauer, B. K. Clark, M. B. Hastings, and M. Troyer. Gate-count estimates for performing quantum chemistry on small quantum computers. *Phys. Rev. A*, **90**, 1 (2014).

[282] H. Weimer, M. Müller, I. Lesanovsky, P. Zoller, and H. P. Büchler. A Rydberg quantum simulator. *Nat. Phys.*, **6**, 382 (2010).

[283] J. Welch, D. Greenbaum, S. Mostame, and A. A. Aspuru-Guzik. Efficient quantum circuits for diagonal unitaries without ancillas. *New J. Phys.*, **16**, 033040 (2014).

[284] J. D. Whitfield. Unified views of quantum simulation algorithms for chemistry. *e-print arXiv: 1502.03771*, pages 1–5 (2015).

[285] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik. Simulation of electronic structure Hamiltonians using quantum computers. *Mol. Phys.*, **109**, 735 (2011).

[286] J. D. Whitfield, P. J. Love, and A. Aspuru-Guzik. Computational complexity in electronic structure. *Phys. Chem. Chem. Phys.*, **15**, 397 (2013).

[287] N. Wiebe, D. W. Berry, P. Hoyer, and B. C. Sanders. Higher Order Decompositions of Ordered Operator Exponentials. *J. Phys. A Math. Theor.*, **43**, 1 (2010).

[288] N. Wiebe, D. W. Berry, P. Hoyer, and B. C. Sanders. Simulating quantum dynamics on a quantum computer. *J. Phys. A Math. Theor.*, **44**, 445308 (2011).

[289] S. Wiesner. Simulations of many-body quantum systems by a quantum computer. *e-print arXiv: 9603028* (1996).

[290] A. J. Williamson, R. Q. Hood, and J. C. Grossman. Linear-scaling quantum Monte Carlo calculations. *Phys. Rev. Lett.*, **87**, 246406 (2001).

[291] Z. Xing and W. Zhang. MaxSolver: An efficient exact algorithm for (weighted) maximum satisfiability. *Artif. Intell.*, **164**, 47 (2005).

[292] J.-S. Xu, M.-H. Yung, X.-Y. Xu, S. Boixo, Z.-W. Zhou, C.-F. Li, A. Aspuru-Guzik, and G.-C. Guo. Demon-like algorithmic quantum cooling and its realization with quantum optics. *Nat. Photon.* (2014).

[293] L. Xue, J. W. Godden, F. L. Stahura, and J. Bajorath. Design and evaluation of a molecular fingerprint involving the transformation of property descriptor values into a binary classification scheme. *J. Chem. Inf. Comput. Sci.*, **43**, 1151 (2003).

[294] D. Yu, X. Huang, H. Wang, Y. Cui, Q. Hu, and R. Zhou. Short-term Solar Flare Level Prediction Using a Bayesian Network Approach. *Astrophys. J.*, **710**, 869 (2010).

[295] K. Yue and K. A. Dill. Forces of tertiary structural organization in globular proteins. *Proc. Natl. Acad. Sci. U. S. A.*, **92**, 146 (1995).

[296] M.-H. Yung, J. Casanova, A. Mezzacapo, J. McClean, L. Lamata, A. Aspuru-Guzik, and E. Solano. From transistor to trapped-ion computers for quantum chemistry. *Sci. Rep.*, **4**, 9 (2014).

[297] M.-H. Yung, J. D. Whitfield, S. Boixo, D. G. Tempel, and A. Aspuru-Guzik. Introduction to Quantum Algorithms for Physics and Chemistry. In *Quantum Inf. Comput. Chem.*, volume 154 of *Advances in Chemical Physics*, pages 67–106. John Wiley & Sons, Inc., 2014.

[298] C. Zalka. Efficient Simulation of Quantum Systems by Quantum Computers. *Fortschritte der Phys.*, **46**, 877 (1998).

[299] Z. Zimborás, M. Faccin, Z. Kádár, J. D. Whitfield, B. P. Lanyon, and J. Biamonte. Quantum Transport Enhancement by Time-Reversal Symmetry Breaking. *Sci. Rep.*, 3 (2013).

[300] M. Ziólkowski, B. Jansík, P. Jorgensen, and J. Olsen. Maximum locality in occupied and virtual orbital spaces using a least-change strategy. *J. Chem. Phys.*, **131**, 124112 (2009).