APPLICATIONS OF NOISY INTERMEDIATE-SCALE QUANTUM COMPUTING TO MANY-BODY NUCLEAR PHYSICS

By

Benjamin Prescott Hall

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

Physics—Doctor of Philosophy
Computational Mathematics, Science and Engineering—Dual Major

2022

ABSTRACT

Many-body nuclear physics is the bridge that takes us from the fundamental laws governing individual nucleons to understanding how groups of them interact to form the nuclei that lie at the heart of all atoms - the building blocks of our universe. Many powerful techniques of classical computation have been developed over the years in order to study ever more complex nuclear systems. However, we seem to be approaching the limits of such classical techniques as the complexity of many-body quantum systems grows exponentially with system size. Yet, the recent development of quantum computers offers one hope as they are predicted to provide a significant advantage over classical computers in certain applications, such as the quantum many-body problem. In this thesis, we focus on developing and applying algorithms to solve various many-body nuclear physics problems that can be run on the near-term quantum computers of the current noisy intermediate-scale quantum (NISQ) era. As these devices have small qubit counts and high noise levels, we focus our algorithms on various many-body toy models in order to gain insight and build a foundation upon which future algorithms will be built to tackle the intractable problems of our time. First, we tailor current quantum algorithms to efficiently run on NISQ era devices and apply them to three pairing models of many-body nuclear physics: the Lipkin model, the Richardson pairing model, and collective neutrino oscillations. We estimate the ground-state energy of the first two models and simulate the time evolution and characterize the entanglement of the third. Then, we develop two novel algorithm to increase the efficiency and applicability of current NISQ era algorithms: an algorithm to compress circuit depth to allow for less noisy computation, and a variational method to prepare an important class of quantum states called Dicke states. Error mitigation techniques used to

improve the accuracy of results are also discussed and employed. All together, this work provides a road map to apply the quantum computers of tomorrow to solve what nuclear phenomena mystify us today.

This thesis is dedicated to my wife, Katherine - my rock,
and to my parents, David and Caroline - my first teachers.

ACKNOWLEDGEMENTS

This dissertation would not exist without the support and guidance of my parents. They sparked a passion for learning in me from a young age and gave me every opportunity necessary to pursue whatever interested me. They have always been there for me. To my wife Katherine, thank you for being my rock. It's so kind of you to always listen to me ramble on about physics and quantum computing. Without your love and support, I would not have made it this far. To my brother Evan, thank you for being there for me to relax with when I needed to recharge. I must also thank my fellow physicist, Jacob Watkins for being such a good friend and making graduate school so enjoyable. Finally, I must thank Pepper, my dog, for getting me through the final push to finish this thesis.

On for the academic side, I would like to thank my advisor Morten Hjorth-Jensen for his guidance and mentorship. He has been my biggest cheerleader throughout my graduate career. He's let me explore a wide range of topics that interest me throughout my graduate studies. Whenever I felt stuck or unworthy, his positivity would pick me right back up. I must also thank all those that mentored me during the various internships I completed during graduate school. To Joseph Carlson, Alessandro Roggero, and Alessandro Baroni, thank you for mentoring me through the publishing of my first paper. I learned a great deal about the fundamentals of quantum computing and neutrino physics from you. And thank you to Patrick Coles for founding and running the summer school at Los Alamos that made that work possible. To Gaute Hagen and Thomas Papenbrock from Oak Ridge, thank you for guiding me through the nuclear theory required to complete the section of my thesis that sparked my interest in this whole topic. I'd like to thank the Department of Energy's Office of Science Graduate Student Research Program for connecting me with

Oak ridge. Next, let me thank Davide Venturelli and everyone at the NASA QuAIL and NAMS team for taking me in and letting me get experience running circuits on a real quantum computer (thank you Riggeti) for the first time. That internship made me realize my passion for pure quantum computing and circuit design. Furthermore, to Alessandro Lavato and Yuri Alexeev at Argonne, thank you for teaching me about classical machine learning and helping me better understand my own research on nuclear pairing. Finally, let me thank all of the secretaries that have helped me along the way, without whom nothing would have ever gotten done: Thank you to Amanda Martinez and Ashlee Adams from LANL, Ryan Smith and Samantha Summers from DOE SCGSR, Tristyn Acasio and Saba Hussain from NASA, and Caity Hitchens and Tamra Lagerwall from Argonne. A special thanks to Kim Crosslan at MSU who's always been there to guide me through my years here and Elizebeth Deliyski at FRIB for helping me navigate the wonderful organization that is nuclear physics at Michigan State.

Finally, let me acknowledge all of the scientists that have come before me and discovered the knowledge upon which this thesis builds.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ALGORITHMS

CHAPTER 1

INTRODUCTION

Classical computation seems to be approaching the limit of its ability to simulate many-body nuclear systems. Such systems are crucial to the understanding of nuclei, the heart of every atom and thus one of the fundamental building blocks of our universe. The Hilbert spaces of such systems grow exponentially with the number of particles in the system. This often necessitates exponential growth of the computational complexity required to solve such systems on a classical computer. Even modern methods, such as quantum Monte Carlo, are hindered by problems, in this case, the fermion sign problem. The recent development of quantum computers promises a plausible path forward to overcoming these hurtles. Fault-tolerant quantum computers with a large enough number of qubits could simulate large many-body quantum systems with ease. For example, Shor's algorithm [75] has been proven to be able to find the eigenenergies of a Hamiltonian with polynomial complexity and time-evolution quantum algorithms benefit from a linear increase in qubits with system size and access to quantum gates that simulate such evolution naturally. However, today's quantum computers are in what has been called the noisy intermediate-scale quantum (NISQ) era, meaning that they have a small number of qubits (order 100) and suffer from substantial noise due to qubit decoherence and measurement errors. Thus, a large part of quantum computing research is currently focused on developing so called near-term algorithms which can run on NISQ era quantum devices. Such algorithms are often hybrid, meaning that they rely on both a quantum and a classical computer, and variational, meaning that they are non-deterministic and rely on optimization algorithms as a sub-routine. One such algorithm, the variational quantum eigensolver, has been used

to solve for the ground state energy of molecules as large as BeH_2 [49] using six qubits and hydrogen chains as large as H_{12}[64] with twelve qubits and 72 two-qubit gates. On the nuclear physics side, the binding energy of the deuteron has been calculated on a quantum computer [28]. In order to scale the applications of such algorithms to larger systems which we cannot currently simulate classically, we will need to develope techniques that allow for short depth circuit implementation on devices with limited qubit connectivity. Furthermore, such algorithms will need to take advantage of specifics of the problem it is simulating such as its symmetries and known approximations. Such is the central aim of this thesis.

In this work, we develope and apply near-term algorithms to toy models of many-body nuclear physics. The algorithms are used to estimate energy eigenvalues of the systems, simulate time-evolution, and characterize entanglement. The techniques developed here serve as a foundation upon which future algorithms can be developed to solve larger systems whose simulation evades current classical computation. This thesis is organized as follows.

In chapters 2 and 3, we provide the background theory upon which the rest of the thesis rests. Chapter 2 gives an overview of many-body nuclear theory, starting with quantum mechanics itself, then going through the formalism used to describe many-body systems. It ends with a review of classical computational techniques that are used to solve many-body quantum problems. Chapter 3 gives an introduction to quantum computing including its basic building blocks: qubits, gates, and circuits. It then gives an explanation of the variational quantum eigensolver (VQE) and an overview of how to map problems from fermionic space to the spin space required for quantum computation.

Chapters 4 through 6 are the heart of the thesis in that they develope techniques and

apply near-term algorithms to three toy pairing models of quantum many-body nuclear physics. Chapter 4 focuses on the Lipkin model. Various classical solutions to the model are discussed before walking through potential ways to solve the model on a quantum computer. While the Lipkin model has been modeled on a quantum computer before, this thesis provides a novel, short-depth ansatz. In chapter 5, we turn our attention to the pairing model, whose classical solutions are discussed and against which quantum solutions are bench-marked. The model is mapped to quantum gates and a novel ansatz is developed for shortening the depth of the quantum circuit on a quantum computer with circular qubit connectivity. A novel algorithm for calculating the energies of excited states is also developed along with a new ansatz for solving the model with a large pairing strength. Various initializations of the ansatz are introduced and compared. In chapter 6, we simulate the time evolution and characterize the entanglement of collective neutrino oscillations. The Hamiltonian is partitioned and the ansatz is constructed in a clever way so as to maximize the quality of the quantum computer's results. Error mitigation techniques are introduced and applied in order to depress the interference of noise in the calculations of the quantum computer.

Next, chapters 7 and 8 introduce two novel quantum algorithms which can be generally applied to various areas of quantum computing in order to improve results. Chapter 7 lays out the quantum circuit squeezing algorithm (QCSA) which trades an increase in number of qubits for a decrease in circuit depth, thus potentially decreasing the accumulation of noise. Chapter 8 discusses a novel approach to prepare Dicke states variationally. Finally, in chapter 9 we present our conclusions and perspective on potential future work in applications of NISQ era algorithms to many-body nuclear physics.

CHAPTER 2

MANY-BODY NUCLEAR THEORY

2.1 Introduction

In this chapter, we introduce the underlying theory of many-body nuclear physics, starting with the basics of quantum mechanics, upon which all else is built. We then give an introduction of the formalisms required to represent many-body systems, including second quantization and particle-hole formalism. Finally, we describe various classical computation techniques that are used to solve such many-body systems.

2.2 Quantum Mechanics

2.2.1 Introduction

To understand many-body nuclear theory, we must first understand the theory underling all of nuclear physics: quantum mechanics. While the theory, like many physical theories, has no exact beginning, one of its formalisms was put forth by the 1920s by Erwin Schrödinger [73]. It is the fundamental theory upon which all other physics is built, setting forth the laws that physical phenomena must follow at the smallest of scales. Quantum mechanics, unlike classical mechanics (the set of laws that physical phenomena follow at large scales) is inherently probabilistic. This means that the equations of quantum mechanics can only give us the probability distribution for the possible values of physical observables not predict, with certainty, what values will be measured. To thoroughly introduce the theory, we begin with the postulates of the theory.

2.2.2 Postulates of Quantum Mechanics

Quantum mechanics is built on a set of postulates from which the rest of the theory is constructed. The postulates are as follows:

Postulate 1: The state of a physical system is represented by a quantum state $|\Psi\rangle$ which belongs to a Hilbert space H.

Postulate 2: Every observable physical quantity $\mathcal{A}$ is represented by a Hermitian operator A acting in H.

Postulate 3: The result of measuring an observable $\mathcal{A}$ must be one of the eigenvalues of its corresponding operator A.

Postulate 4: When an observable $\mathcal{A}$ is measured in the state $|\Psi\rangle$, the probability of obtaining the eigenvalue a_n of A is given by the squared amplitude of its corresponding eigen-vector $|\Phi_n\rangle$. One can expand the state $|\Psi\rangle$ in terms of $|\Phi_n\rangle$, the eigenvectors of A, as follows

$$|\Psi\rangle = \sum_n c_n |\Phi_n\rangle . \tag{2.1}$$

Thus $P(a_n)$, the probability of obtaining the eigenvalue a_n can be determined by

$$P(a_n) = |c_n|^2 = |\langle\Phi_n|\Psi\rangle|^2. \tag{2.2}$$

Postulate 5: If the measurement of an observable $\mathcal{A}$ results in a_n then the state collapses to the normalized projection of $|\Psi\rangle$ onto the eigen-subspace associated with a_n

$$\Psi \rightarrow \frac{P_n |\Psi\rangle}{\sqrt{\langle\Psi|P_n|\Psi\rangle}}, \tag{2.3}$$

where

$$P_n = \sum_m |\Phi_m\rangle \langle\Phi_m| , \tag{2.4}$$

and m sums over the eigenvectors that correspond to the eigenvector a_n.

Postulate 6: The state $|\Psi\rangle$ evolves in time according to the Schrödinger equation

$$i\hbar\frac{d}{dt}|\Psi(t)\rangle = H|\Psi(t)\rangle\,. \tag{2.5}$$

2.2.3 The Schrödinger Equation

It is often useful to think about the time-independent and time-dependent parts of a system separately. To separate the Schrödinger equation into time-dependent and time-independent components, we assume that the state $|\Psi(x,t)\rangle$ can be written as the product of a time-independent $|X(x)\rangle$ and a time-dependent $|T(t)\rangle$ state:

$$|\Psi(x,t)\rangle = |X(x)\rangle\,|T(t)\rangle\,. \tag{2.6}$$

The sixth postulate of quantum mechanics tells us that the state $|\Psi\rangle$ of a system evolves in time according to the Schrödinger equation (2.5). Plugging the separation (2.6) into the Schrödinger equation (2.5) allows us to separate the differential equation:

$$i\hbar\frac{d}{dt}|\Psi(x,t)\rangle = H|\Psi(x,t)\rangle\,, \tag{2.7}$$

becomes

$$i\hbar\frac{d}{dt}|X(x)\rangle\,|T(t)\rangle = H|X(x)\rangle\,|T(t)\rangle\,, \tag{2.8}$$

which allows

$$\frac{1}{|T(t)\rangle}i\hbar\frac{d}{dt}|T(t)\rangle = \frac{1}{|X(x)\rangle}H|X(x)\rangle = E, \tag{2.9}$$

where we were able to set both sides of (2.9) equal to the constant E because they each exclusively depend on different parameters, namely position x an time t. From this, we arrive at the time-independent Schrödinger equation (2.10) and time-evolution equation (2.11) given below

$$H\left|X(x)\right\rangle = E\left|X(x)\right\rangle, \tag{2.10}$$

$$i\hbar\frac{d}{dt}\left|T(t)\right\rangle = E\left|T(t)\right\rangle. \tag{2.11}$$

Solving the time evolution equation (2.11) yields the equation for the time-dependent state

$$\left|T(t)\right\rangle = e^{-iEt}, \tag{2.12}$$

where we've set $\hbar = 1$. This implies that the full state is given by

$$\begin{aligned}\left|\Psi(x,t)\right\rangle &= e^{-iEt}X(x)\\ &= \sum_{n=0}^{\infty}\frac{(-it)^n}{n!}E^nX(x)\\ &= \sum_{n=0}^{\infty}\frac{(-it)^n}{n!}H^nX(x)\\ &= U(t)X(x),\end{aligned} \tag{2.13}$$

where we've defined the time-evolution operator $U(t)$ as

$$U(t) = e^{-iHt}. \tag{2.14}$$

Here we've both assumed that H is time-independent and used the time-independent Schrödinger equation (2.10).

2.3 Many-Body Formalism

2.3.1 Introduction

Systems with multiple particles in quantum mechanics are called many-body systems. The formalism describing such systems was developed most notably by Vladimir Fock [34]. In this formalism, which is called second-quantization, one keeps track of the number of particles contained in each state. It is a powerful tool for understanding large systems of particles in quantum mechanics.

2.3.2 Single-particle states

We define the coordinate representation of a single-particle state

$$\psi_p(x) = \langle x|p\rangle\,, \tag{2.15}$$

to mean that the particle labeled p is occupying the state which is characterized by the set of coordinates x. Consider a complete set of orthonormal, single-particle states $P = \{p_1, \ldots, p_n\}$. The orthogonality of the set means that for any states $p, q \in P$, their inner product is equal to the Kronecker delta,

$$\delta_{p,q} = \langle p|q\rangle = \int dx\, \langle p|x\rangle\, \langle x|q\rangle = \int dx \psi_p^*(x)\psi_q(x). \tag{2.16}$$

The completeness of the set means that any single particle state can be expanded as a linear combination of the set

$$|\psi\rangle = \sum_p c_p\, |p\rangle\,, \tag{2.17}$$

where

$$c_p = \langle p|\psi\rangle = \int dx \psi_p^*(x)\psi(x). \tag{2.18}$$

2.3.3 Multi-particle states

We now wish to consider a system of multiple identical fermions, each of which lives in a Hilbert space H. Such a multi-particle systems lives in the Fock space $F(H)$ which is the direct sum of the anti-symmetric tensor product of the single-particle Hilbert spaces H,

$$F(H) = \bigoplus_{N=0}^{\infty} S_- H^{\otimes N}, \tag{2.19}$$

where S_- anti-symmetrizes tensors. Each tensor product of single-particle Hilbert spaces $H^{\otimes N} = \underbrace{H \otimes \ldots \otimes H}_{N \text{ times}}$ contains multi-particle states

$$\bigotimes_{i=1}^{N} \psi_{p_i}(x_i) = \langle x_1 \cdots x_N | p_1 \cdots p_N \rangle = | p_1 \cdots p_N \rangle \,, \tag{2.20}$$

where $|p_1 \cdots p_n\rangle = \bigotimes_{i=1}^{N} |p_i\rangle$. As indicated above, $\langle x_1 \cdots x_N|$ is often dropped, and assumed to always be in order $(1, \ldots, N)$. Thus, if particles are permuted according to a permutation matrix σ, the state $|p_{\sigma_1} \cdots p_{\sigma_N}\rangle = \langle x_1 \cdots x_N | p_{\sigma_1} \cdots p_{\sigma_N}\rangle$ implies that particle labeled p_{σ_i} is in the state characterized by the set of coordinates x_i. The spaces $H^{\otimes N}$ are anti-symmetrized because, according to the spin-statistics theorem, fermionic wave-functions must be anti-symmetric with respect to the exchange of particles; that is

$$\left| \cdots p_j \cdots p_i \cdots \right\rangle = - \left| \cdots p_i \cdots p_j \cdots \right\rangle , \tag{2.21}$$

which says that if particles p_i and p_j swap states, the corresponding many-body state picks up a minus sign. In general, a multi-particle fermionic state can pick up a factor of a negative one depending on the permutation of the particles. That is

$$\left| p_{\sigma_1} \ldots p_{\sigma_N} \right\rangle = (-1)^{\sigma} \, | p_1 \cdots p_N \rangle \,, \tag{2.22}$$

where

$$(-1)^{\sigma} = \begin{cases} +1 & \text{if } \sigma \text{ is an even permutation} \\ -1 & \text{if } \sigma \text{ is an odd permutation.} \end{cases} \tag{2.23}$$

Thus, each anti-symmetric tensor product of single-particle Hilbert spaces $S_{-}H^{\otimes N}$ contains anti-symmetrized multi-particle states

$$|\{p_1 \dots p_N\}\rangle = \sqrt{N!} A \, |p_1 \dots p_N\rangle \,, \tag{2.24}$$

where A is called the anti-symmetrizer and is defined as

$$A = \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^{\sigma} \sigma, \tag{2.25}$$

and S_N is the symmetric group of order N. Applying the definition of the anti-symmetrizer (2.25) to the definition of the anti-symmetrized multi-particle states (2.24) yields

$$|\{p_1 \dots p_N\}\rangle = \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{\sigma} \sigma \, |p_1 \dots p_N\rangle \tag{2.26}$$

$$= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{\sigma} \, |p_{\sigma_1} \dots p_{\sigma_N}\rangle \tag{2.27}$$

$$= \frac{1}{\sqrt{N!}} |P|, \tag{2.28}$$

where P is an $N \times N$ matrix whose entries are equal to $P_{ij} = \psi_{p_j}(x_i)$; that is

$$|\{p_1 \cdots p_N\}\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_{p_1}(x_1) & \psi_{p_2}(x_1) & \cdots & \psi_{p_N}(x_1) \\ \psi_{p_1}(x_2) & \psi_{p_2}(x_2) & \cdots & \psi_{p_N}(x_2) \\ \vdots & \vdots & & \vdots \\ \psi_{p_1}(x_N) & \psi_{p_2}(x_N) & \cdots & \psi_{p_N}(x_N) \end{vmatrix}. \tag{2.29}$$

The above determinant is called a Slater determinant [77]. One can see that this state is anti-symmetric with respect to the exchange of two particles as this corresponds to exchanging two columns of the determinant, which picks up a factor of negative one. One can see that the set of anti-symmetrized multi-particle states is orthonormal as, for any two multi-particle state $|\{p_1 \cdots p_n\}\rangle, |\{q_1 \cdots q_n\}\rangle \in S_- H^{\otimes N}$ their overlap is

$$\langle\{p_1 \cdots p_n\}|\{q_1 \cdots q_n\}\rangle = \frac{1}{N!}|P^\dagger||Q| = \frac{1}{N!}|P^\dagger Q| = \begin{cases} 1 & \text{if } p_i = q_i \text{ for all } i = 1, \ldots, N \\ 0 & \text{otherwise,} \end{cases} \tag{2.30}$$

as the entries of $P^\dagger Q$ are given by

$$(P^\dagger Q)_{ij} = \sum_k^N P^\dagger_{ik} Q_{kj} = \sum_k^N \psi^*_{p_i}(x_k)\psi_{q_j}(x_k) = \langle p_i|q_j\rangle = \delta_{p_i q_j}, \tag{2.31}$$

which implies

$$|P^\dagger Q| = \sum_{\sigma \in S_N} (-1)^\sigma \prod_{i=1}^N \delta_{p_i q_{\sigma_i}}. \tag{2.32}$$

Additionally, the set of anti-symmetrized multi-particle states is also complete. This means that any anti-symmetrized multi-particle state can be expanded as

$$|\psi\rangle = \sum_{p_1,\ldots,p_N} c_{p_1,\ldots,p_N} |\{p_1 \cdots p_N\}\rangle, \tag{2.33}$$

where the coefficients $c_{p_1,\ldots,p_N}$ are given by

$$c_{p_1,\ldots,p_N} = \langle\{p_1 \cdots p_N\}|\psi\rangle = \int d_{x_1} \ldots d_{x_n} \psi^*_{p_1}(x_1) \ldots \psi^*_{p_N}(x_N)\psi(x_1, \ldots, x_N). \tag{2.34}$$

2.3.4 Second Quantization

Second quantization (occupation number representation) is a formalism in which multi-particle systems are described with the information of how many particles are in each state.

Its central ideas where first introduced by in 1927 Paul Dirac [24]. A fermionic multi-particle state is described in second quantization as

$$|\{p_1 \dots p_N\}\rangle = |\dots n_{p_1} \dots n_{p_N} \dots\rangle\,, \tag{2.35}$$

where

$$n_i = \begin{cases} 1 & \text{if } p_i \in \{p_1, \dots, p_N\} \\ 0 & \text{if } p_i \notin \{p_1, \dots, p_N\}. \end{cases}$$

Note that the occupation numbers n_i can only be 0 or 1 as, according to the Pauli exclusion principle, each state can contain at most one fermion. In this representation, it is convenient to introduce creation ($a^\dagger$) and annihilation (a) operators which add and remove particles from states, respectively. Acting on a fermionic single-particle state, they behave as follows:

$$\begin{aligned} a^\dagger |0\rangle &= |1\rangle\,, & a^\dagger |1\rangle &= 0, \\ a_i |0\rangle &= 0, & a_i |1\rangle &= |0\rangle\,. \end{aligned} \tag{2.36}$$

It can be seen that trying to increase the number of particles in a state above 1 or decrease the number of particles in a state below 0 results in 0, reflecting the Pauli-exclusion principle. Their behavior on multi-fermionic states is as follows:

$$a_i^\dagger |n_1 \dots n_n\rangle = (-1)^{N_i}(1 - n_i) |n_1 \dots n_{i-1} 1 n_{i+1} \dots n_n\rangle\,, \tag{2.37}$$

$$a_i |n_1 \dots n_n\rangle = (-1)^{N_i} n_i |n_1 \dots n_{i-1} 0 n_{i+1} \dots n_n\rangle\,, \tag{2.38}$$

where $N_i = \sum_{j=1}^{i-1} n_j$. The $(-1)^{N_i}$ factor comes from the anti-commutation relations of the fermionic operators

$$\left\{a_i, a_j^\dagger\right\} = \delta_{ij},$$

$$\left\{a_i, a_j\right\} = \{a_i^\dagger, a_j^\dagger\} = 0, \tag{2.39}$$

which, in turn, come from the fact that exchanging fermions picks up a factor of negative one. Note that these operators can be used to create arbitrary fermionic multi-particle states from the vacuum $|0\rangle$ as follows:

$$\prod_{i=1}^{N} a_{p_i}^\dagger |0\rangle = |\{p_1 \cdots p_N\}\rangle \,. \tag{2.40}$$

One popular combination of these operators is the number operator

$$N_p = a_p^\dagger a_p, \tag{2.41}$$

which counts the number of particles in a state (p),

$$N_p \left|\cdots n_p \cdots\right\rangle = n_p \left|\cdots n_p \cdots\right\rangle . \tag{2.42}$$

These operators can also be used to represent arbitrary k-body operators in second quantization as follows:

$$O_k = \left(\frac{1}{k!}\right)^2 \sum_{\substack{p_1 \ldots p_k \\ q_1 \ldots q_k}} \langle p_1 \ldots p_k | o_k | q_1 \ldots q_k \rangle_{\mathrm{A}} \, a_{p_1}^\dagger \ldots a_{p_k}^\dagger a_{q_k} \ldots a_{q_1}, \tag{2.43}$$

where the anti-symmetrized k-body matrix element is defined as

$$\langle p_1 \ldots p_k | o_k | q_1 \ldots q_k \rangle_A = \langle \{p_1 \ldots p_k\} | o_k | q_1 \ldots q_k \rangle = \langle p_1 \ldots p_k | o_k | \{q_1 \ldots q_k\} \rangle \,, \tag{2.44}$$

where

$$\begin{aligned} &\langle p_1 \ldots p_k | o_k | q_1 \ldots q_k \rangle \\ &= \int dx_1 \ldots dx_k \psi_{p_1}^*(x_1) \ldots \psi_{p_k}^*(x_k) o_k(x_1, \ldots, x_k) \psi_{q_1}(x_1) \ldots \psi_{q_k}(x_k). \end{aligned} \tag{2.45}$$

Note that from here on out in this thesis we will drop the brackets in Dirac-ket notation when referring to anti-symmetrized states as we will be dealing exclusively with fermions (whose states are always anti-symmetric). That is, from now on all $|x\rangle$ should be taken to mean $|\{x\}\rangle$.

2.3.5 Particle-Hole Formalism

When dealing with a large number of particles, it can become convenient to introduce a reference state ($|\Phi_0\rangle$) other than the vacuum ($|0\rangle$). To do this, we define a Fermi level (F) below which all states are assumed to be occupied and above which, all states are assumed to be unoccupied. We call the states below the Fermi level hole states and those above, particle states. The hole states are indexed with $i, j, k, ...$ while the particle states are indexed with $a, b, c, ...$ If we wish to remain ambiguous as to whether or not a state is a particle or hole state, we index with $p, q, r, ...$ Thus, our reference state

$$|\Phi_0\rangle = |\{i_1 i_2 \dots i_n\}\rangle\,, \tag{2.46}$$

in this case with n hole states, acts as the new vacuum state. Because our new vacuum already contains some particles, the notions of creation and annihilation must be redefined. This is because, for example, annihilation of a particle that exists in the reference state vacuum does not go to zero like it would if acting upon the true vacuum. Therefore, we introduce particle-hole creation and annihilation operators

$$b_p^\dagger = \begin{cases} a_p^\dagger & \text{if } p > F \\ a_p & \text{if } p \leq F \end{cases}, \qquad b_p = \begin{cases} b_p & \text{if } p > F \\ b_p^\dagger & \text{if } p \leq F \end{cases}. \tag{2.47}$$

In this way, applying a particle-hole creation operator (to the reference state) for a particle in the reference state can be thought of as creating a hole (when, in reference to the true

vacuum, it destroys a particle), and vise verse.

2.3.6 Contractions

When calculating quantities with a large number of particles, it can become tedious to manipulate the expression via successive applications of the fermionic anti-commutation relations (2.39). One tool to alleviate this is the contraction. The contraction of two operators X and Y is defined as

$$\overbrace{XY} = XY - N(XY), \tag{2.48}$$

where $N(XY)$ indicates the normal order of XY. The normal order of a string of operators is the one that vanishes when its expectation value is taken in the vacuum: $\langle 0| N(XY) |0\rangle = 0$. The normal order of pairs of the fermionic operators are

$$N(a_p a_q) = a_p a_q, \tag{2.49}$$

$$N(a_p^\dagger a_q^\dagger) = a_p^\dagger a_q^\dagger, \tag{2.50}$$

$$N(a_p^\dagger a_q) = a_p^\dagger a_q, \tag{2.51}$$

$$N(a_p a_q^\dagger) = -a_q^\dagger a_p. \tag{2.52}$$

It follows that the contractions of the fermionic operators are

$$\overbrace{a_p a_q} = 0, \tag{2.53}$$

$$\overbrace{a_p^\dagger a_q^\dagger} = 0, \tag{2.54}$$

$$\overbrace{a_p^\dagger a_q} = 0, \tag{2.55}$$

$$\overbrace{a_p a_q^\dagger} = \delta_{pq}. \tag{2.56}$$

In the particle-hole formalism, our normal orderings change as we have redefined the vacuum: $N(XY) = \langle\Phi_0| XY |\Phi_0\rangle$. The normal order of pairs of particle-hole operators are

$$N(a_a a_b) = a_a a_b, \qquad N(a_i a_j) = a_i a_j, \tag{2.57}$$

$$N(a_a^\dagger a_b^\dagger) = a_a a_b, \qquad N(a_i^\dagger a_j^\dagger) = a_i a_j, \tag{2.58}$$

$$N(a_a^\dagger a_b) = a_a^\dagger a_b, \qquad N(a_i^\dagger a_j) = -a_j^\dagger a_i, \tag{2.59}$$

$$N(a_a a_b^\dagger) = -a_b^\dagger a_a, \qquad N(a_i a_j^\dagger) = a_i^\dagger a_j. \tag{2.60}$$

It follows that the contractions of the fermionic operators are

$$\overbracket{a_a a_b} = 0, \qquad \overbracket{a_i a_j} = 0, \tag{2.61}$$

$$\overbracket{a_a^\dagger a_b^\dagger} = 0, \qquad \overbracket{a_i^\dagger a_j^\dagger} = 0, \tag{2.62}$$

$$\overbracket{a_a^\dagger a_b} = 0, \qquad \overbracket{a_i^\dagger a_j} = \delta_{ij}, \tag{2.63}$$

$$\overbracket{a_a a_b^\dagger} = \delta_{ab}, \qquad \overbracket{a_i a_j^\dagger} = 0. \tag{2.64}$$

One important theorem involving contractions is Wick's theorem [86] which states that a string of operators $ABCD\ldots$ can be written as a sum of normal-ordered strings involving all possible contractions; that is

$$ABCD\ldots = N(ABCD\ldots) + \sum_{\text{singles}} N(\overbracket{AB}CD\ldots) + \sum_{\text{doubles}} N(\overbracket{AB}\overbracket{CD}\ldots) + \ldots, \tag{2.65}$$

where the sums over singles and doubles refer to all the possible ways the string of operators can be contracted once and twice, respectively. Wick's theorem is especially useful for calculating expectation values as only the fully-contracted terms in the sum survive (all other terms have leftover normal order terms which, by definition, have an expectation value of zero). For example

$$\langle 0| ABCD |0\rangle = \langle 0| \overbracket{AB}\overbracket{CD} |0\rangle + \langle 0| \overbracket{A\overbracket{BC}D} |0\rangle . \tag{2.66}$$

2.4 Classical Computation Techniques

2.4.1 Full Configuration Interaction

Consider the time-independent Schrodinger equation

$$H |\psi\rangle = E |\psi\rangle . \tag{2.67}$$

Expanding $|\psi\rangle$ it in terms of an orthonormal basis $\{|n\rangle\}$

$$|\psi\rangle = \sum_n c_n |n\rangle , \tag{2.68}$$

and inserting it into the Schrodinger equation (2.67), yields

$$\sum_n c_n H |n\rangle = E \sum_n c_n |n\rangle . \tag{2.69}$$

Inserting the identity $\sum_k |k\rangle \langle k|$ into the left-hand side yields

$$\sum_n H_{kn} c_n |k\rangle = E \sum_n c_n |n\rangle , \tag{2.70}$$

where $H_{kn} = \langle k|H|n\rangle$. Left-multiplying both sides by $\langle m|$ yields

$$\sum_n H_{mn} c_n = E c_m, \tag{2.71}$$

which can be written in matrix form

$$HC = EC. \tag{2.72}$$

Thus, diagonalizing H will yield the eigenvalues E and corresponding eigenfunctions which are constructed from the coefficients c_n according to (2.68).

2.4.2 Symmetry Method

While full-configuration interaction theory can, in theory, be applied to any problem, the trade-off is that, in order for this to be true, FCI is a problem agnostic theory. That is, it does not take advantage of any of the information about the Hamiltonian itself. Without this information, one is left to diagonalize a $2^{2\Omega} \times 2^{2\Omega}$ matrix whose size grows exponentially with the number of states Ω. However, if one incorporates some information about the Hamiltonian, such as its symmetries, then one will only have to diagonalize several, small matrices. The reason why will be seen in the following explanation of the symmetry method.

In the symmetry method, one starts by identifying the symmetries S_k of the Hamiltonian, elements of the symmetry groups S that commute with the Hamiltonian

$$[S_k, H] = 0. \tag{2.73}$$

The states are labeled by the quantum numbers corresponding to the irreducible representations of the symmetry groups. The Hamiltonian matrix is a block-diagonal matrix, with one block for each irreducible representation. The proof is as follows: Let S_k be an element of a symmetric group S which commutes with the Hamiltonian. Also, let $S_k |\psi_n\rangle = s_n |\psi_n\rangle$. Then

$$\langle\psi_n|H|\psi_m\rangle = \frac{1}{s_m} \langle\psi_n|HS_k|\psi_m\rangle \tag{2.74}$$

$$= \frac{1}{s_m} \langle\psi_n|S_kH|\psi_m\rangle \tag{2.75}$$

$$= \frac{s_n}{s_m} \langle\psi_n|H|\psi_m\rangle\,, \tag{2.76}$$

which implies that $(s_m - s_n)H_{nm} = 0$. Thus, either $s_n = s_m$ or $H_{nm} = 0$ and $s_n \neq s_m$ ($|\psi\rangle_n$ and $|\psi\rangle_m$ don't have the same eigenvalue with respect to S_k). Thus, the Hamiltonian matrix

is block diagonal.

2.4.3 Hartree-Fock Theory

Hartree-Fock theory attempts to approximate the ground state of a system by treating it as a collection of non-interacting particles subject to a mean-field potential that approximates their interaction. In the theory, which was finalized in 1935 by Hartree [40], one varies the single-particle orbital basis of a single Slater-determinant $|\Phi\rangle$ in order to minimize the energy of that state; that is, the Hartree-Fock energy E_{HF} is defined as

$$E_{\mathrm{HF}} = \min_{|\Phi\rangle} \frac{\langle\Phi|H|\Phi\rangle}{\langle\Phi|\Phi\rangle}. \tag{2.77}$$

The variational principle guarantees that $E_{\mathrm{HF}} \geq E_0$ where E_0 is the ground state energy of H. Minimizing the functional will yield a set of equations which can be solved to determine the Hartree-Fock Hamiltonian and its eigenvalues (single-particle energies) [55]. A small variation in $|\Phi\rangle$

$$|\Phi\rangle \rightarrow |\Phi\rangle + |\delta\Phi\rangle\,, \tag{2.78}$$

leads to the following variation in the energy

$$\frac{\langle\Phi|H|\Phi\rangle}{\langle\Phi|\Phi\rangle} \rightarrow \frac{\langle\Phi|H|\Phi+\delta\Phi\rangle}{\langle\Phi|\Phi+\delta\Phi\rangle} \tag{2.79}$$

$$= \frac{\langle\Phi|H|\Phi\rangle}{\langle\Phi|\Phi\rangle + \langle\Phi|\delta\Phi\rangle} + \frac{\langle\Phi|H|\delta\Phi\rangle}{\langle\Phi|\Phi\rangle + \langle\Phi|\delta\Phi\rangle} \tag{2.80}$$

$$= \frac{\langle\Phi|H|\Phi\rangle}{\langle\Phi|\Phi\rangle}\left(1 - \frac{\langle\Phi|\delta\Phi\rangle}{\langle\Phi|\Phi\rangle}\right) + \frac{\langle\Phi|H|\delta\Phi\rangle}{\langle\Phi|\Phi\rangle} + O\left(\langle\Phi|\delta\Phi\rangle^2\right) \tag{2.81}$$

$$= E + \delta E, \tag{2.82}$$

up to first order in $\langle\Phi|\delta\Phi\rangle$, where

$$\delta E = \langle\delta\Phi|H - E|\Phi\rangle\,, \tag{2.83}$$

since $H|\Phi\rangle = E|\Phi\rangle$ and $\langle\Phi|\Phi\rangle = 1$. The energy is stationary when

$$0 = \delta E = \langle\delta\Phi|H - E|\Phi\rangle\,. \tag{2.84}$$

To make it so, we must first determine the form of $|\delta\Phi\rangle$. With

$$|\Phi\rangle = \prod_{i=1}^{n} a_i^\dagger\,|0\rangle\,, \tag{2.85}$$

infinitesimally varying the basis

$$a_i^\dagger \to \widetilde{a}_i^\dagger = \sum_p (\delta_{ip} + \epsilon_{ip}) a_p^\dagger, \tag{2.86}$$

leads to the following variation in the wavefunction

$$\begin{aligned}
|\widetilde{\Phi}\rangle &= \prod_{i=1}^{n} \widetilde{a}_i^\dagger\,|0\rangle && (2.87)\\
&= \prod_{i=1}^{n} \left[\sum_p (\delta_{ip} + \epsilon_{ip}) a_p^\dagger\right]|0\rangle && (2.88)\\
&= \prod_{i=1}^{n} \left[a_i^\dagger + \sum_p \epsilon_{ip} a_p^\dagger\right]|0\rangle && (2.89)\\
&= \prod_{i=1}^{n} a_i^\dagger\,|0\rangle + \sum_p \left[\left(\prod_{j=i+1}^{n} a_j^\dagger\right)\epsilon_{ip} a_p^\dagger \left(\prod_{j=1}^{i-1} a_j^\dagger\right)\right]|0\rangle + \mathcal{O}(\epsilon^2) && (2.90)\\
&= |\Phi\rangle + |\delta\Phi\rangle\,, && (2.91)
\end{aligned}$$

up to first order in ϵ where

$$|\delta\Phi\rangle = \sum_p \left[\left(\prod_{j=i+1}^{n} a_j^\dagger\right)\epsilon_{ip} a_p^\dagger \left(\prod_{j=1}^{i-1} a_j^\dagger\right)\right]|0\rangle\,, \tag{2.92}$$

which we can rewrite as

$$|\delta\Phi\rangle = \sum_a \left[\left(\prod_{j=i+1}^{n} a_j^\dagger\right)\epsilon_{ia} a_a^\dagger a_i a_i^\dagger \left(\prod_{j=1}^{i-1} a_j^\dagger\right)\right]|0\rangle \tag{2.93}$$

$$= \sum_a \epsilon_{ia} a_a^\dagger a_i \left[\left(\prod_{j=i+1}^{n} a_j^\dagger \right) a_i^\dagger \left(\prod_{j=1}^{i-1} a_j^\dagger \right) \right] |0\rangle \tag{2.94}$$

$$= \sum_a \epsilon_{ia} a_a^\dagger a_i \, |\Phi\rangle \,, \tag{2.95}$$

since p must be a particle state in order for $|\delta\Phi\rangle \neq 0$ and

$$\begin{aligned} a_i a_i^\dagger \left(\prod_{j=1}^{i-1} a_j^\dagger \right) |0\rangle &= (\{a_i, a_i^\dagger\} - a_i^\dagger a_i) \left(\prod_{j=1}^{i-1} a_j^\dagger \right) |0\rangle \\ &= (1 - a_i^\dagger) \left(\prod_{j=1}^{i-1} a_j^\dagger \right) a_i \, |0\rangle \\ &= \left(\prod_{j=1}^{i-1} a_j^\dagger \right) |0\rangle \,. \end{aligned} \tag{2.96}$$

Having determined the form of $|\delta\Phi\rangle$, we can plug it into the condition that the energy is stationary (2.84) to yield

$$0 = \langle \delta\Phi | H - E | \Phi \rangle \tag{2.97}$$

$$= \sum_a \epsilon_{ia} \langle \Phi | \, a_i^\dagger a_a (H - E) \, | \Phi \rangle \tag{2.98}$$

$$= \sum_a \epsilon_{ia} \left(\langle \Phi | \, a_i^\dagger a_a H \, | \Phi \rangle - E \cancel{\langle \Phi | \, a_i^\dagger a_a \, | \Phi \rangle} \right) , \tag{2.99}$$

which holds if

$$0 = \langle \Phi | \, a_i^\dagger a_a H \, | \Phi \rangle \tag{2.100}$$

$$= \sum_{pq} t_q^p \langle \Phi | \, a_i^\dagger a_a a_p^\dagger a_q \, | \Phi \rangle + \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} \langle \Phi | \, a_i^\dagger a_a a_p^\dagger a_q^\dagger a_s a_r \, | \Phi \rangle \tag{2.101}$$

$$= \sum_{pq} t_q^p \langle \Phi | \, \overbrace{a_i^\dagger a_a a_p^\dagger} a_q \, | \Phi \rangle$$

$$+ \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} \left[\langle \Phi | \, \overbrace{a_i^\dagger a_a a_p^\dagger a_q^\dagger a_s} a_r \, | \Phi \rangle + \langle \Phi | \, \overbrace{a_i^\dagger a_a a_p^\dagger a_q^\dagger a_s a_r} \, | \Phi \rangle \right.$$

$$+ \langle \Phi | \, a_i^\dagger a_a a_p^\dagger a_q^\dagger a_s a_r \, | \Phi \rangle + \langle \Phi | \, a_i^\dagger a_a a_p^\dagger a_q^\dagger a_s a_r \, | \Phi \rangle \Bigg] \tag{2.102}$$

$$= \sum_{pq} t_q^p \delta_{pa} \delta_{qi} + \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} (h(s) \delta_{pa} \delta_{qs} \delta_{ri} - h(s) \delta_{ps} \delta_{qa} \delta_{ri}$$

$$+ \; h(r) \delta_{pa} \delta_{qr} \delta_{si} - h(r) \delta_{pr} \delta_{qa} \delta_{si}) \tag{2.103}$$

$$= t_i^a + \frac{1}{4} \sum_j \left(v_{ij}^{aj} - v_{ij}^{ja} - v_{ji}^{aj} + v_{ji}^{ja} \right) \tag{2.104}$$

$$= t_i^a + \sum_j v_{ij}^{aj}, \tag{2.105}$$

which can be written as $\langle a | H_{\mathrm{HF}} | i \rangle = 0$ by defining the Hartree-Fock Hamiltonian

$$H_{\mathrm{HF}} = t + \sum_{rsi} v_{qi}^{pj} \, |r\rangle \langle s| \, . \tag{2.106}$$

The single-particle energies are the eigenvalues of the Hartree-Fock Hamiltonian

$$\epsilon_p = \langle p | \, H_{\mathrm{HF}} \, | p \rangle \tag{2.107}$$

$$= t_p^p + \sum_i v_{pi}^{pi}, \tag{2.108}$$

where we've relabeled $j \leftrightarrow i$.

2.4.4 Many-Body Perturbation Theory

To go beyond Hartree-Fock, we introduce many-body permutation theory [52] which adds in correlations between particles as perturbations to the Hartree-Fock wavefunction. Analogous to single-body perturbation theory, it is possible to approximate the energy of a system by writing the Hamiltonian as the sum of two terms, one of which is readily diagonalizable. An introduction of this theory is given here as its results will be used to inform our initialization of the variational quantum algorithms which will be used later. Our aim is to solve the Schrödinger equation

$$H \, |\Psi_0\rangle = E \, |\Psi_0\rangle \, . \tag{2.109}$$

For perturbation theory, we assume that our Hamiltonian H can be written as the sum of an unperturbed Hamiltonian H_0 and an interacting Hamiltonian H_I

$$H = H_0 + H_I, \tag{2.110}$$

where the Schrödinger equation for H_0

$$H_0 \left|\Phi_n\right\rangle = E_n \left|\Phi_n\right\rangle, \tag{2.111}$$

is easily solvable. We expand the exact wave-function $|\Psi_0\rangle$ in terms of the unperturbed wave-function

$$\left|\Psi_0\right\rangle = \left|\Phi_0\right\rangle + \sum_{n=1}^{\infty} C_n \left|\Phi_n\right\rangle. \tag{2.112}$$

Employing intermediate normalization $\langle\Phi_0|\Psi_0\rangle = 1$, along with equations (2.109) through (2.111), we can derive

$$E = \left\langle\Phi_0|H|\Psi_0\right\rangle = \left\langle\Phi_0|H_0|\Psi_0\right\rangle + \left\langle\Phi_0|H_I|\Psi_0\right\rangle = E_0 + \left\langle\Phi_0|H_I|\Psi_0\right\rangle, \tag{2.113}$$

from which we define

$$\Delta E = E - E_0 = \left\langle\Phi_0|H_I|\Psi_0\right\rangle. \tag{2.114}$$

We now introduce the following operators

$$P = \left|\Phi_0\right\rangle\left\langle\Phi_0\right|, \tag{2.115}$$

$$Q = \sum_{n=1}^{\infty} \left|\Phi_n\right\rangle\left\langle\Phi_n\right|. \tag{2.116}$$

Noting that $P + Q = 1$ and using the expanded form of the exact wave-function (2.112) we can rewrite the exact wave-function as

$$\left|\Psi_0\right\rangle = (P + Q)\left|\Psi_0\right\rangle = \left|\Phi_0\right\rangle + Q\left|\Psi_0\right\rangle. \tag{2.117}$$

To determine the second term, note that by plugging in the summed version of the Hamiltonian (2.110) into the exact Schrödinger equation (2.109), rearranging, and adding the term $E_0 |\Psi_0\rangle$ (for Rayleigh-Schrodinger perturbation theory) to both sides, one arrives at

$$(E_0 - H_0) |\Psi_0\rangle = (H_I - \Delta E) |\Psi_0\rangle , \tag{2.118}$$

where we've used the definition of ΔE (2.114). Applying Q to both sides implies

$$Q |\Psi_0\rangle = \frac{Q}{E_0 - H_0}(H_I - \Delta E) |\Psi_0\rangle . \tag{2.119}$$

Plugging this into (2.117) yields

$$|\Psi_0\rangle = |\Phi_0\rangle + \frac{Q}{E_0 - H_0}(H_I - \Delta E) |\Psi_0\rangle , \tag{2.120}$$

Inserting this equation into itself iteratively yields

$$|\Psi_0\rangle = \sum_{n=0}^{\infty} \left[\frac{Q}{E_0 - H_0}(H_I - \Delta E) \right]^n |\Phi_0\rangle , \tag{2.121}$$

which, when plugged into the expression for the energy (2.114), yields

$$\Delta E = \sum_{n=0}^{\infty} \langle \Phi_0 | H_I \left[\frac{Q}{E_0 - H_0}(H_I - \Delta E) \right]^n |\Phi_0\rangle . \tag{2.122}$$

Note that since Q commutes with H_0 and ΔE is constant, we have that $Q\Delta E |\Phi_0\rangle = \Delta E Q |\Phi_0\rangle = 0$ and thus the energy becomes

$$\Delta E = \langle \Phi_0 | H_I | \Phi_0 \rangle + \sum_{n=1}^{\infty} \langle \Phi_0 | H_I \left[\frac{Q}{E_0 - H_0}(H_I - \Delta E) \right]^{n-1} \frac{Q}{E_0 - H_0} H_I |\Phi_0\rangle . \tag{2.123}$$

Perturbatively expanding in term of H_I, which is assumed to be small, we can write

$$\Delta E = \sum_{n=1}^{\infty} \Delta E^{(n)} . \tag{2.124}$$

Comparing (2.123) and (2.124), we identify

$$\Delta E^{(1)} = \langle \Phi_0 | H_I | \Phi_0 \rangle \,, \tag{2.125}$$

$$\Delta E^{(2)} = \langle \Phi_0 | H_I \frac{Q}{E_0 - H_0} H_I | \Phi_0 \rangle \,, \tag{2.126}$$

$$\Delta E^{(3)} = \langle \Phi_0 | H_I \frac{Q}{E_0 - H_0} H_I \frac{Q}{E_0 - H_0} H_I | \Phi_0 \rangle \,, \tag{2.127}$$

$$- \langle \Phi_0 | H_I \frac{Q}{E_0 - H_0} \langle \Phi_0 | H_I | \Phi_0 \rangle \frac{Q}{E_0 - H_0} H_I | \Phi_0 \rangle \,, \tag{2.128}$$

where we've used (2.114) for the last expression. Recall that we assumed that our Hamiltonian

$$H = \sum_{pq} t_q^p a_p^\dagger a_q + \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r, \tag{2.129}$$

could be partitioned (2.110) as $H = H_0 + H_I$ where H_0 is easily solvable. There are several ways to partition the Hamiltonian as such. The one we will consider here is called Hartree-Fock partitioning [80] in which we set the unperturbed Hamiltonian H_0 equal to the Hartree-Fock Hamiltonian (2.106) since it can easily be diagonalized as

$$H_0 = \sum_p \epsilon_p a_p^\dagger a_p, \tag{2.130}$$

where ϵ_p are the single-particle energies (2.107) which are given by $\epsilon_p = t_p^p + \sum_i v_{pi}^{pi}$. With this partitioning, the interacting Hamiltonian H_I becomes

$$\begin{aligned} H_I &= H - H_0 \\ &= \sum_{pq} (t_q^p - \epsilon_p \delta_{pq}) a_p^\dagger a_q + \frac{1}{4} \sum_{pqrs} v_{rs}^{pq} a_p^\dagger a_q^\dagger a_s a_r. \end{aligned} \tag{2.131}$$

Applying the unperturbed Hamiltonian H_0 (2.130) to the reference state $|\Phi_0\rangle$ yields

$$H_0 |\Phi_0\rangle = \sum_p \epsilon_p a_p^\dagger a_p \left(\prod_n a_{q_n}^\dagger \right) |0\rangle \tag{2.132}$$

$$= \sum_p \epsilon_p \sum_m \left[a_p^\dagger \overbrace{a_p \left(\prod_{n=1}^{m-1} a_{q_n}^\dagger \right)} a_{q_m}^\dagger \left(\prod_{n=m+1}^{N} a_{q_n}^\dagger \right) \right] |0\rangle \tag{2.133}$$

$$= \sum_p \epsilon_p \sum_m (-1)^{m-1} \delta_{pq_m} a_p^\dagger \left(\prod_n a_{q_n}^\dagger \right) |0\rangle \tag{2.134}$$

$$= \sum_p \epsilon_p \sum_m \delta_{pq_m} \left(\prod_n a_{q_n}^\dagger \right) |0\rangle \tag{2.135}$$

$$= \sum_m \epsilon_{q_m} |\Phi_0\rangle \,, \tag{2.136}$$

which implies that, with the notation $H_0 |\Phi_0\rangle = E_0 |\Phi_0\rangle$, we have

$$E_0 = \sum_n \epsilon_n, \tag{2.137}$$

where we've relabeled $m \to n$. Using (2.136), we can see that applying H_0 to excited forms of the reference state yields

$$E_{i_1 \ldots i_n}^{a_1 \ldots a_n} \left| \Phi_{i_1 \ldots i_n}^{a_1 \ldots a_n} \right\rangle = H_0 \left| \Phi_{i_1 \ldots i_n}^{a_1 \ldots a_n} \right\rangle \tag{2.138}$$

$$= \left(\sum_{q_n \notin \{i_1, \ldots, i_n\}} \epsilon_{q_n} + \sum_{m=1}^{n} \epsilon_{a_m} \right) |0\rangle \tag{2.139}$$

$$= \left(\sum_{q_n} \epsilon_{q_n} + \sum_{m=1}^{n} \epsilon_{a_m} - \sum_{m=1}^{n} \epsilon_{i_m} \right) |0\rangle \tag{2.140}$$

$$= \left(E_0 - \epsilon_{i_1 \ldots i_n}^{a_1 \ldots a_n} \right) |0\rangle \,, \tag{2.141}$$

which implies that

$$E_0 - E_{i_1 \ldots i_n}^{a_1 \ldots a_n} = \epsilon_{i_1 \ldots i_n}^{a_1 \ldots a_n}, \tag{2.142}$$

where we've defined

$$\epsilon_{i_1 \ldots i_n}^{a_1 \ldots a_n} = \sum_{m=1}^{n} \left(\epsilon_{i_m} - \epsilon_{a_m} \right). \tag{2.143}$$

With Hartree-Fock partitioning, we can use (2.142) to write the correlation energies (2.125 - 2.127) more explicitly. Starting with $\Delta E^{(2)}$ (2.126) we have

$$\Delta E^{(2)} = \langle \Phi_0 | \, H_I \frac{Q}{E_0 - H_0} H_I \, | \Phi_0 \rangle \,, \tag{2.144}$$

and noting that

$$\frac{Q}{E_0 - H_0} = Q \, (E_0 - H_0)^{-1} \, Q \tag{2.145}$$

$$= \sum_{n,m=1} |\Phi_n\rangle \, \langle \Phi_n | \, (E_0 - H_0)^{-1} \, |\Phi_m\rangle \, \langle \Phi_m | \tag{2.146}$$

$$= \sum_{n,m=1} \delta_{nm} \, |\Phi_n\rangle \, (E_0 - E_m)^{-1} \, \langle \Phi_m | \tag{2.147}$$

$$= \sum_{n=1} |\Phi_n\rangle \, (E_0 - E_n)^{-1} \, \langle \Phi_n | \tag{2.148}$$

since Q is idempotent, we have that

$$\sum_{ai} \frac{\langle \Phi_0 | H_I | \Phi_i^a \rangle \, \langle \Phi_i^a | H_I | \Phi_0 \rangle}{\epsilon_i^a} + \sum_{a<b,i<j} \frac{\langle \Phi_0 | H_I | \Phi_{ij}^{ab} \rangle \, \langle \Phi_{ij}^{ab} | H_I | \Phi_0 \rangle}{\epsilon_{ij}^{ab}}, \tag{2.149}$$

as any higher terms would be automatically zero as H only includes up to two-body interactions. The only non-zero contribution is

$$\langle \Phi_{ij}^{ab} | H_I | \Phi_0 \rangle = \sum_{pq} v_{rs}^{pq} \, \langle \Phi_0 | \, \{a_i^\dagger a_j^\dagger a_a a_b\} \{a_p^\dagger a_q^\dagger a_s a_r\} \, | \Phi_0 \rangle \tag{2.150}$$

$$= \frac{1}{4} \sum_{pq} v_{rs}^{pq} \, \langle \Phi_0 | \, \{\overbrace{a_i^\dagger a_j^\dagger a_a a_b}\} \{\overbrace{a_p^\dagger a_q^\dagger a_s a_r}\} \, | \Phi_0 \rangle$$

$$+ \langle \Phi_0 | \, \{\overbrace{a_i^\dagger a_j^\dagger a_a a_b}\} \{\overbrace{a_p^\dagger a_q^\dagger a_s a_r}\} \, | \Phi_0 \rangle$$

$$+ \langle \Phi_0 | \, \{\overbrace{a_i^\dagger a_j^\dagger a_a a_b}\} \{\overbrace{a_p^\dagger a_q^\dagger a_s a_r}\} \, | \Phi_0 \rangle$$

$$+ \langle \Phi_0 | \{ \overbrace{a_i^\dagger a_j^\dagger a_a a_b}\} \{ \overbrace{a_p^\dagger a_q^\dagger a_s a_r}\} | \Phi_0 \rangle \tag{2.151}$$

$$= \frac{1}{4} \sum_{pq} v_{rs}^{pq} (\delta_{ir}\delta_{js}\delta_{aq}\delta_{bp} - \delta_{is}\delta_{jr}\delta_{aq}\delta_{bp}$$

$$- \delta_{ir}\delta_{js}\delta_{ap}\delta_{bq} + \delta_{is}\delta_{jr}\delta_{ap}\delta_{bq}) \tag{2.152}$$

$$= \frac{1}{4} \left(v_{ij}^{ba} - v_{ji}^{ba} - v_{ij}^{ab} + v_{ji}^{ab} \right) \tag{2.153}$$

$$= v_{ji}^{ab}. \tag{2.154}$$

In an analogous manner, one can compute $\langle \Phi_0 | H_I | \Phi_{ij}^{ab} \rangle = v_{ab}^{ji}$ which yields

$$\Delta E^{(2)} = \sum_{a<b, i<j} \frac{v_{ij}^{ab} v_{ab}^{ij}}{\epsilon_{ij}^{ab}}. \tag{2.155}$$

The same procedure can be carried out for higher order terms [53]. For example, the third order contribution to the correlation energy is given by

$$\Delta E^{(3)} = \sum_{\substack{i<j<k \\ a<b<c}} \frac{v_{ab}^{ij} v_{ic}^{bk} v_{ik}^{ac}}{\epsilon_{ij}^{ab} \epsilon_{ik}^{ac}} + \sum_{\substack{i<j \\ a<b<c<d}} \frac{v_{cd}^{ij} v_{ab}^{cd} v_{ij}^{ab}}{\epsilon_{ij}^{ab} \epsilon_{ij}^{cd}} + \sum_{\substack{i<j<k<l \\ a<b}} \frac{v_{kl}^{ab} v_{ij}^{kl} v_{ab}^{ij}}{\epsilon_{ij}^{ab} \epsilon_{kl}^{ab}}. \tag{2.156}$$

2.4.5 Coupled Cluster Theory

Coupled cluster theory (CC) was first developed to study nuclear physics by Coester and Kümmel [32] in the 1950's. The ansatz for the theory is given by

$$|\Psi\rangle = e^T |\Phi_0\rangle, \tag{2.157}$$

where $|\Phi_0\rangle$ is the reference state and T, the cluster operator, is defined as

$$T = \sum_{p=n}^{A} T_n, \tag{2.158}$$

where A is the maximum number of particle-hole excitations and each term in the sum is given by

$$T_n = \frac{1}{n!} \sum_{\substack{i_1 \ldots i_n \\ a_1 \ldots a_n}} t_{i_1 \ldots i_n}^{a_1 \ldots a_n} a_{a_1}^\dagger \ldots a_{a_n}^\dagger a_{i_n} \ldots a_{i_1}. \tag{2.159}$$

Often the cluster operator is truncated to a small number of terms. A common example is to truncate to $N = 2$ leading to $T = T_1 + T_2$; this is the so-called singles and doubles approximation (CCSD). Starting with the time-independent Schrödinger equation

$$H |\Psi\rangle = E |\Psi\rangle , \tag{2.160}$$

inserting the coupled cluster ansatz (2.157) and left-multiplying both sides by e^{-T}, and then left-multiplying by either the reference state or an excited state yields

$$E = \langle \Phi_0 | \overline{H} | \Phi_0 \rangle , \tag{2.161}$$

$$0 = \langle \Phi_{i_1 \ldots i_n}^{a_1 \ldots a_n} | \overline{H} | \Phi_0 \rangle , \tag{2.162}$$

respectively. Here $\bar{H}$ is the similarity transformed Hamiltonian

$$\overline{H} = e^{-T} H e^T. \tag{2.163}$$

Additionally, the excited states notation is as follows

$$|\Phi_{i_1 \ldots i_n}^{a_1 \ldots a_n}\rangle = a_{a_1}^\dagger \ldots a_{a_n}^\dagger a_{i_1} \ldots a_{i_n} |\Phi_0\rangle . \tag{2.164}$$

Finally, after defining the reference energy

$$E_{\text{ref}} = \langle \Phi | H | \Phi \rangle , \tag{2.165}$$

we define the correlation energy (ΔE) to be the difference between the CC energy E as defined in 2.161 and the reference energy

$$\Delta E = E - \langle \Phi | H | \Phi \rangle \,, \tag{2.166}$$

which subtracts of the reference energy. Note that the energy (2.161) depends on the cluster amplitudes $t_{i_1 \ldots i_n}^{a_1 \ldots a_n}$, which can be obtained deterministically by solving the amplitude equations (2.162). In order to do so, one often expands the similarity transformed Hamiltonian $\overline{H}$ (2.163) using the Baker-Campbell-Hausdorff (BCH) identity

$$\begin{aligned} e^{-T} H e^{T} &= e^{\mathrm{ad}_T} H \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \mathrm{ad}_T^n H \\ &= H + [H, T] + \frac{1}{2} [[H, T], T] + \cdots \,, \end{aligned} \tag{2.167}$$

where $\mathrm{ad}_a b = [b, a]$. When one plugs this expansion into energy expression (2.161) it can be shown that, when the reference state $|\Phi_0\rangle$ is a single determinant, the expansion terminates at fourth order ($n = 4$). Coupled cluster with single determinant reference states usually performs well at equilibrium configurations but poorly for strongly correlated systems. This shortcoming can be overcome with unitary coupled cluster theory.

2.4.6 Pair Coupled Cluster Doubles Theory

In pair coupled cluster doubles theory [42], often abbreviated as pCCD, the cluster operator (2.158) is set to a variant of T_2 called T_p which is restricted to moving pairs of fermions. That is, $T = T_p$, where

$$T_p = \sum_{ia} t_i^a A_a^\dagger A_i. \tag{2.168}$$

Here, $A^\dagger$ and A are the pair fermionic creation and annihilation operators, respectively, defined as

$$A_p^\dagger = a_{p+}^\dagger a_{p-}^\dagger \tag{2.169}$$

$$A_p = a_{p-} a_{p+}, \tag{2.170}$$

where $p+$ and $p-$ index the spin-up and spin-down fermions in the p^{th} energy level, respectively. The coupled cluster equations (2.161 and 2.162) become

$$E = \langle \Phi_0 | \overline{H} | \Phi_0 \rangle \tag{2.171}$$

$$0 = \left\langle \Phi_i^a \middle| \overline{H} \middle| \Phi_0 \right\rangle, \tag{2.172}$$

where the similarity transformed Hamiltonian $\overline{H}$ is

$$\overline{H} = e^{-T_p} H e^{T_p}, \tag{2.173}$$

and the excited state $\langle \Phi_i^a |$ is obtained by

$$\langle \Phi_i^a | = \langle \Phi_0 | \, A_i^\dagger A_a. \tag{2.174}$$

Note that we truncated the amplitude equation (2.162) to a single excited state for the pair coupled cluster doubles amplitude equations.

2.4.7 Unitary Coupled Cluster Theory

In unitary coupled cluster (UCC) theory [8], the cluster operator T is replaced with the purely imaginary operator $T - T^\dagger$, leading to the unitary exponential ansatz

$$|\Psi\rangle = e^{T - T^\dagger} |\Phi_0\rangle. \tag{2.175}$$

This ansatz is variational and thus the ground state energy E_0 can be obtained from the variational principle

$$E_0 = \min_t \langle \Phi_0 | \overline{H} | \Phi_0 \rangle \,, \tag{2.176}$$

where t are the cluster amplitudes, of which T is a function, and $\overline{H}$ is the similarity transformed Hamiltonian

$$\overline{H} = e^{T^\dagger - T} H e^{T - T^\dagger}. \tag{2.177}$$

While unitary coupled cluster theory overcomes some weaknesses of coupled cluster theory, it is not classically tractable to implement. One can see this by expanding, analogously to (2.167), the similarity transformed Hamiltonian for UCC using the BCH identity,

$$\begin{aligned} e^{T^\dagger - T} H e^{T - T^\dagger} =& e^{\mathrm{ad}_{T - T^\dagger}} H & (2.178) \\ =& \sum_{n=0}^{\infty} \frac{1}{n!} \mathrm{ad}^n_{T - T^\dagger} H \\ =& H + [H, T] - [H, T^\dagger] + \frac{1}{2}([[H, T], T] + [[H, T^\dagger], T^\dagger] \\ & - [[H, T], T^\dagger] - [[H, T^\dagger], T]) + \cdots , & (2.179) \end{aligned}$$

which, unlike the expansion for CC, does not naturally truncate. Thus, it is classically intractable to solve the resulting amplitude and energy equations. Fortunately, unitary operators are implementable on quantum computers (indeed, as all quantum gates must be unitary). Thus, UCC is a viable candidate ansatz for the so-called Variational Quantum Eigensolver (VQE), a quantum algorithm which is discussed in the following chapter.

CHAPTER 3

QUANTUM COMPUTING

3.1 Introduction

The idea of quantum computing was first put forth in 1982 by Richard Feynman [33]. The initial motivation was to simulate quantum mechanics itself because, as Feynman said "Nature isn't classical, dammit, and if you want to make a simulation of nature, you'd better make it quantum mechanical..." Quantum computers use qubits, rather than bits, as their fundamental units. A qubit is an abstraction of a quantum binary event. The states qubits are manipulated by physical processes whose abstractions are referred to as quantum gates. Today's quantum computers have of order 100 qubits and can implement of order 1000 gates [19] before the accrued noise becomes insurmountable.

3.2 Qubits

The bit is the basic unit of information in classical computing. It is the abstraction of a binary logical state. One denotes the two possible states as 0 and 1. That is, a bit c is

$$c = 0 \text{ OR } 1. \tag{3.1}$$

A bit can be realized physically by a classical binary event, a classical event that has exactly two distinct outcomes. Examples include the toss of a coin, the presence or absence of a hole in a paper card, or (as used in modern computers) two levels of electric charge stored in a capacitor. Multiple bits are written together as a bit-string

$$c = c_0 c_2 \ldots c_{n-1}, \tag{3.2}$$

where $c_i = 0, 1$ for $i = 0, 1, \ldots, n-1$. Thus, n qubits can represent 2^n distinct logical states.

The qubit, by contrast, is the basic unit of information in quantum computing. It is the abstraction of a quantum binary event. Examples include an electron being spin-up or spin-down or a quantum system being in one of two energy states. The state of a qubit is spanned by the computational basis which consists of the following two quantum states

$$|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{3.3}$$

$$|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{3.4}$$

Thus, a qubit is a two-level quantum state which can be written as a superposition of the computational basis states, $|0\rangle$ and $|1\rangle$. That is, a qubit $|q\rangle$ can be written as

$$|q\rangle = a\,|0\rangle + b\,|1\rangle\,, \tag{3.5}$$

where $a, b \in \mathbb{C}$, under the restriction that

$$|a|^2 + |b|^2 = 1. \tag{3.6}$$

One convenient way to represent the state of a single qubit $|q\rangle$ is on the Bloch sphere [4]. To do so, one parameterizes the state's coefficients with two angles θ and ϕ as

$$|\psi\rangle = \cos\frac{\theta}{2}\,|0\rangle + \sin\frac{\theta}{2}e^{i\phi}\,|1\rangle\,. \tag{3.7}$$

Then, the qubit $|\psi\rangle$ can be represented on the Bloch sphere as a vector with polar angle θ and azimuthal angle ϕ as depicted in Figure 3.1. Note that the coefficient of the $|0\rangle$ term in 3.7 does note contain a complex phase. This is because the global phase of a qubit's state is irrelevant and thus the relative phase of the $|0\rangle$ state can always be absorbed into such a global phase.

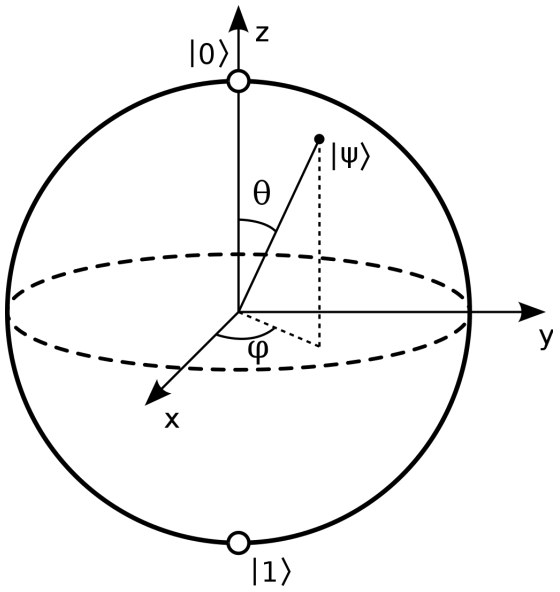


Figure 3.1: Depiction of the state $|\psi\rangle$ on the Bloch sphere [20].

As two by one, complex-valued vectors, qubits can be “multiplied" in several important ways. To explain these manipulations, we must first introduce the adjoint of a qubit. Given a qubit

$$|q\rangle = \begin{pmatrix} a \\ b \end{pmatrix}, \tag{3.8}$$

its adjoint $\langle q|$ is given by its Hermitian conjugate

$$\langle q| = |q\rangle^{\dagger} = \begin{pmatrix} a^* & b^* \end{pmatrix}. \tag{3.9}$$

The inner-product of two qubits $|q_1\rangle$ and $|q_2\rangle$ is defined as

$$\langle q_1|q_2\rangle = \begin{pmatrix} a_1^* & b_1^* \end{pmatrix} \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = a_1^* a_2 + b_1^* b_2. \tag{3.10}$$

The outer-product of two qubits $|q_1\rangle$ and $|q_2\rangle$ is defined as

$$|q_1\rangle \langle q_2| = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \begin{pmatrix} a_2^* & b_2^* \end{pmatrix} = \begin{pmatrix} a_1 a_2^* & a_1 b_2^* \\ b_1 a_2^* & b_1 b_2^* \end{pmatrix}. \tag{3.11}$$

Finally, the tensor-product of two-qubits is defined as

$$|q_1 q_2\rangle = |q_1\rangle \otimes |q_2\rangle = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \otimes \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \begin{pmatrix} a_1 \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \\ b_1 \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} a_1 a_2 \\ a_1 b_2 \\ b_1 a_2 \\ b_1 b_2 \end{pmatrix}. \tag{3.12}$$

3.3 Quantum Gates

Quantum gates are abstractions of the physical actions applied to the physical systems representing qubits. Mathematically, they are complex-valued, unitary matrices which act on the complex-valued, normalized vectors that represent qubits. As the quantum analog of classical logic gates (such as AND and OR), there is a corresponding quantum gate for every classical gate; however, there are quantum gates which have no classical counterpart. They act on a set of qubits, changing the qubit's state in the process. That is, if U is a quantum gate and $|q\rangle$ is a qubit, then acting the gate U on the qubit $|q\rangle$ transforms the qubit as follows:

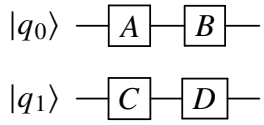
$$|q\rangle \xrightarrow{U} U|q\rangle\,. \tag{3.13}$$

This action can be represented via the following quantum circuit

$$|q\rangle \quad \text{---}\boxed{U}\text{---} \quad U|q\rangle \tag{3.14}$$

Quantum circuits are diagrammatic representations of quantum algorithms. The horizontal dimension corresponds to time; moving left to right corresponds to forward motion in time. They consist of a set of qubits $|q_n\rangle$ which are stacked vertically on the left-hand side of the diagram. Lines, called quantum wires, extend horizontally to the right from each qubit, representing its state moving forward in time. Additionally, they contain a set of quantum gates that interrupt to the quantum wires, implying that they are applied to the corresponding qubit. Gates are applied chronologically, left to right. With this, we can see that the quantum circuit above (3.15) implies that the quantum gate U is being applied to a qubit in state $|q\rangle$.

To explain what quantum circuits represent mathematically, consider the following circuit



(3.15)

which implies the following mathematical statement

$$|q_0 q_1\rangle \rightarrow (B \otimes D)(A \otimes C)\,|q_0 q_1\rangle \tag{3.16}$$

$$\rightarrow (BA) \otimes (DC)\,|q_0 q_1\rangle \tag{3.17}$$

$$\rightarrow BA\,|q_0\rangle\, DC\,|q_1\rangle\,. \tag{3.18}$$

Note that mathematical form is in reverse order from circuit form ($BA \leftrightarrow AB$). This is because the operator closest to the state mathematically (furthest to the right on a quantum circuit) acts first. Additionally, we are able to write the actions of the top two and bottom two gates as acting separately on each qubit as every gate here is a single-qubit gate (acting on only one qubit). The same would not be true for certain two-qubit gates which entangle

the states of the two qubits, not allowing their state to be written in a separable form. Finally, we define the depth of a quantum circuit as the number of columns of gates. The circuit above thus has a depth of two because it contains two columns of gates, namely $A \otimes C$ and $B \otimes D$.

3.3.1 Single-Qubit Gates

A single-qubit gate is the abstraction of a physical action that is applied to one qubit. It can be represented by a matrix $U \in \mathrm{SU}(2)$. Any single-qubit gate can be parameterized by three angles: θ, ϕ, and λ as follows

$$U(\theta, \phi, \lambda) = \begin{pmatrix} \cos\frac{\theta}{2} & -e^{i\lambda}\sin\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} & e^{i(\phi+\lambda)}\cos\frac{\theta}{2} \end{pmatrix}. \tag{3.19}$$

A common set of single-qubit gates is the set of Pauli gates, whose members correspond to the Pauli matrices

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{3.20}$$

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \tag{3.21}$$

$$Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \tag{3.22}$$

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \tag{3.23}$$

which satisfy the relation

$$[\sigma, \tau] = i\epsilon_{\sigma\tau\upsilon}\upsilon, \tag{3.24}$$

for $\sigma, \tau, \upsilon \in \{X, Y, Z\}$. These gates form a basis for the algebra $\mathfrak{su}(2)$. Exponentiating them will thus give us a basis for SU(2), the group within which all single-qubit gates live. These exponentiated Pauli gates are called rotation gates $R_\sigma(\theta)$ because they rotate the quantum state around the axis $\sigma = X, Y, Z$ of the Bloch sphere (Figure 3.1) by an angle θ. They are defined as

$$R_X(\theta) = e^{-i\frac{\theta}{2}X} = \begin{pmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \tag{3.25}$$

$$R_Y(\theta) = e^{-i\frac{\theta}{2}Y} = \begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}, \tag{3.26}$$

$$R_Z(\theta) = e^{-i\frac{\theta}{2}Z} = \begin{pmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{pmatrix}. \tag{3.27}$$

Because they form a basis for SU(2), any single-qubit gate can be decomposed into three rotation gates. Indeed

$$R_z(\phi)R_y(\theta)R_z(\lambda) = \begin{pmatrix} e^{-i\phi/2} & 0 \\ 0 & e^{i\phi/2} \end{pmatrix}\begin{pmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{pmatrix}\begin{pmatrix} e^{-i\lambda/2} & 0 \\ 0 & e^{i\lambda/2} \end{pmatrix} \tag{3.28}$$

$$= e^{-i(\phi+\lambda)/2}\begin{pmatrix} \cos\frac{\theta}{2} & -e^{i\lambda}\sin\frac{\theta}{2} \\ e^{i\phi}\sin\frac{\theta}{2} & e^{i(\phi+\lambda)}\cos\frac{\theta}{2} \end{pmatrix}, \tag{3.29}$$

which is, up to a global phase, equal to the expression for an arbitrary single-qubit gate (3.19).

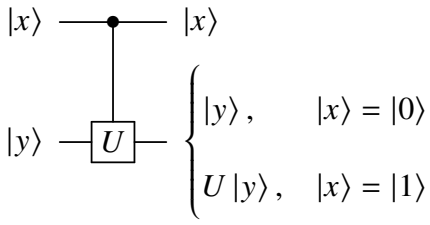
3.3.2 Two-Qubit Gates

A two-qubit gate is the abstraction of a physical action that is applied to two qubits. It can be represented by a matrix U from the group SU(4). One important class of two-qubit

gates are the controlled gates, which work as follows: Suppose U is a single-qubit gate. A controlled-U gate (CU) acts on two qubits: a control qubit $|x\rangle$ and a target qubit $|y\rangle$. The controlled-U gate applies the identity I or the single-qubit gate U to the target qubit if the control gate is in the zero state $|0\rangle$ or the one state $|1\rangle$, respectively. The control qubit is acted upon by the identity I. This can be represented as follows:

$$CU\,|xy\rangle = \begin{cases} |xy\rangle & \text{if } |x\rangle = |0\rangle \\ |x\rangle\, U\,|y\rangle & \text{if } |x\rangle = |1\rangle \end{cases}. \tag{3.30}$$

The action of a controlled-U gate (CU) can be represented via quantum circuit as follows



(3.31)

It can be written in matrix form by writing it as a superposition of the two possible cases, each written as a simple tensor product

$$CU = |0\rangle\langle 0| \otimes I + |1\rangle\langle 1| \otimes U \tag{3.32}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & u_{00} & u_{01} \\ 0 & 0 & u_{10} & u_{11} \end{pmatrix}. \tag{3.33}$$

where u_{ij} (for $i, j \in 0, 1$) are the matrix elements of U. One of the most fundamental controlled gates is the CNOT gate. It is defined as the controlled-X gate CX and thus flips

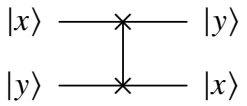
the state of the target qubit if the control qubit is in the zero state $|0\rangle$. It can be written in matrix form as follows:

$$\text{CNOT} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{3.34}$$

A widely used two-qubit gate that goes beyond the simple controlled function is the SWAP gate. It swaps the states of the two qubits that it acts upon

$$\text{SWAP}\,|xy\rangle = |yx\rangle\,, \tag{3.35}$$

as depicted in the quantum circuit below

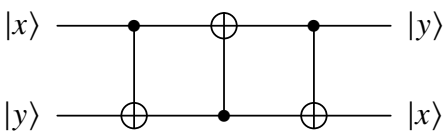


(3.36)

and has the following matrix form

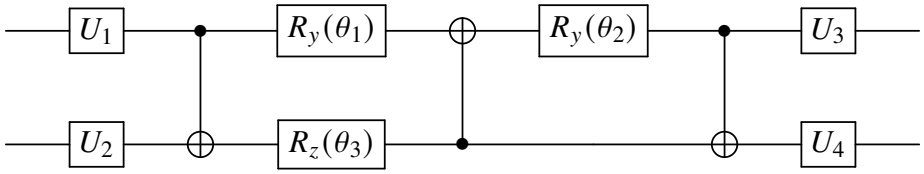
$$\text{SWAP} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{3.37}$$

It can be decomposed into a series of three CNOTs, each of which has its directionality flipped from the previous



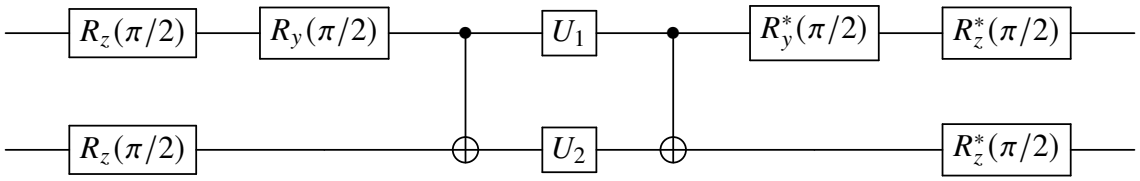
(3.38)

As for arbitrary two-qubit gates $U \in \mathrm{SU}(4)$, they can be optimally decomposed (up to a global phase) into the following sequence [83] fifteen elementary one-qubit gates and three CNOT gates



(3.39)

where U_1, U_2, U_3, U_4 are single-qubit gates, each of which requires three parameters as each can be decomposed into three elementary one-qubit gates (rotation gates). Additionally, the parameters $\theta_1, \theta_2, \theta_3$ are determined by the arbitrary two-qubit gate to be decomposed. Two-qubit gates that are restricted to $U \in \mathrm{SO}(4)$ can be decomposed into a shorter depth circuit consisting of just twelve elementary single-qubit gates and two CNOT gates



(3.40)

3.4 Variational Quantum Eigensolver

3.4.1 Introduction

One of the first algorithms developed to estimate the eigenenergies of a Hamiltonian was quantum phase estimation [50]. In the algorithm, one encodes the eigenenergies, one binary bit at a time (up to n bits), into the complex phases of the quantum states of the Hilbert space for n qubits. It does this by applying powers of controlled unitary evolution operators to a quantum state that can be expanded in terms of the Hamiltonian's eigenvectors. The

eigenenergies are encoded into the complex phases in such a way that taking the inverse quantum Fourier transformation of the states into which the eigenenergies are encoded results in a measurement probability distribution that has peaks around the bit strings that represent a binary fraction which corresponds to the eigenenergies of the quantum state acted upon by the controlled unitary operators. While quantum phase estimation (QPE) is provably efficient, non-hybrid, and non-variational, the number of qubits and circuit length required to execute it is too great for our NISQ era quantum computers. Thus, QPE will only be efficiently applicable on large, fault-tolerant quantum computers that likely won't exist in the near future.

Therefore, a different algorithm for finding the eigenenergies of a quantum Hamiltonian was put forth in 2014 called the variational quantum eigensolver [63], commonly referred to as VQE. The algorithm is hybrid, meaning that it requires the use of both a quantum computer and a classical computer. It is also variational, meaning that it relies, ultimately, on solving an optimization problem by varying parameters and thus is not deterministic like QPE. The variational quantum eigensolver is based on the variational principle: The expectation value of a Hamiltonian H in a state $|\psi(\theta)\rangle$ parameterized by a set of angles θ, is always greater than or equal to the minimum eigenenergy E_0. To see this, let $|n\rangle$ be the eigenstates of H

$$H|n\rangle = E_n|n\rangle. \tag{3.41}$$

We can then expand our state $|\psi(\theta)\rangle$ in terms of said eigenstates

$$|\psi(\theta)\rangle = \sum_n c_n|n\rangle, \tag{3.42}$$

and take the expectation value of the Hamiltonian in this state to yield

$$\begin{aligned}
\langle\psi(\theta)|H|\psi(\theta)\rangle &= \sum_{nm} c_m^* c_n \langle m|H|n\rangle \\
&= \sum_{nm} c_m^* c_n E_n \langle m|n\rangle \\
&= \sum_{nm} \delta_{nm} c_m^* c_n E_n \\
&= \sum_{n} |c_n|^2 E_n \\
&\geq E_0 \sum_{n} |c_n|^2 \\
&= E_0, \qquad (3.43)
\end{aligned}$$

which implies that we can minimize over the set of angles θ and arrive at the ground state energy E_0:

$$\min_{\theta} \langle\psi(\theta)|H|\psi(\theta)\rangle = E_0. \qquad (3.44)$$

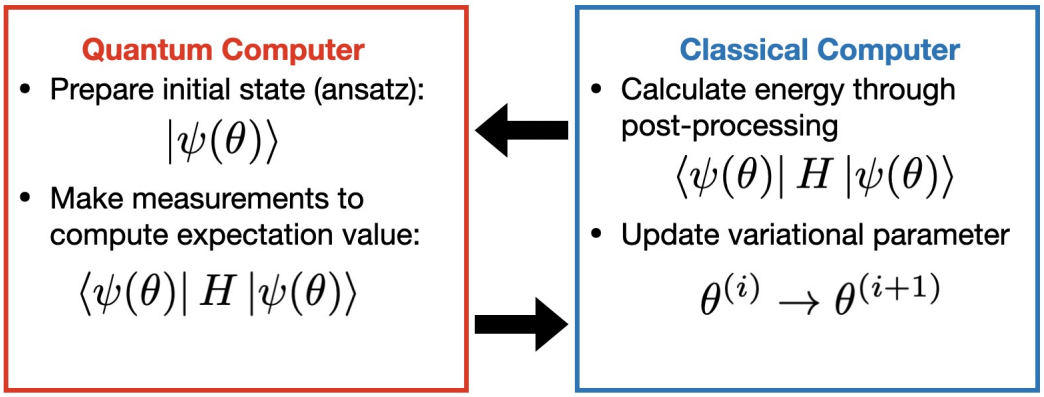


Figure 3.2: Schematic of the Variational Quantum Eigensolver.

Using this fact, VQE can be broken down into the following steps, as noted in Figure 3.2:

1. Prepare the variational state $|\psi(\theta)\rangle$ on a quantum computer.

2. Measure this circuit in various bases and send these measurements to a classical computer.

3. Post-processes the measurement data on the classical computer to compute the expectation value $\langle\psi(\theta)|H|\psi(\theta)\rangle$

4. Vary the parameters θ according to a classical minimization algorithm and send them back to the quantum computer which runs step 1 again.

This loop continues until the classical optimization algorithm terminates which results in a set of angles $\theta_{\min}$ that characterize the ground state $|\phi(\theta_{\min})\rangle$ and an estimate for the ground state energy $E_0 = \langle\psi(\theta_{\min})|H|\psi(\theta_{\min})\rangle$.

3.4.2 Expectation Values

To execute the second step of VQE, we need to understand how expectation values of operators can be estimated via quantum computers by post-processing measurements of quantum circuits in different basis. To rotate bases, one uses the basis rotator B_σ which is defined for each Pauli gate σ to be

$$B_\sigma = \begin{cases} H, & \text{if } \sigma = X \\ HS^\dagger, & \text{if } \sigma = Y \\ I, & \text{if } \sigma = Z \end{cases} . \tag{3.45}$$

Note the following identity of the basis rotator

$$B_\sigma^\dagger Z B_\sigma = \sigma, \tag{3.46}$$

which follows from the fact that $HZH = X$ and $SXS^\dagger = Y$. With this, we see that the expectation value of an arbitrary Puali-gate σ in the state $|\psi\rangle$ can be expressed as a linear combination of probabilities

$$\begin{aligned} E_\psi(\sigma) &= \langle\psi|\sigma|\psi\rangle \\ &= \langle\psi|B_\sigma^\dagger Z B_\sigma|\psi\rangle \\ &= \langle\phi|Z|\phi\rangle \\ &= \langle\phi|\left(\sum_{x\in\{0,1\}} (-1)^x |x\rangle\langle x|\right)|\phi\rangle \\ &= \sum_{x\in\{0,1\}} (-1)^x |\langle x|\phi\rangle|^2 \\ &= \sum_{x\in\{0,1\}} (-1)^x P(|\phi\rangle \to |x\rangle), \end{aligned} \tag{3.47}$$

where $|\phi\rangle = |B_\sigma\phi\rangle$ and $P(|\phi\rangle \to |x\rangle)$ is the probability that the state $|\phi\rangle$ collapses to the state $|x\rangle$ when measured. This can be extended to any arbitrary Pauli string as follows: consider the string of Pauli operators $P = \bigotimes_{p\in Q} \sigma_p$ which acts non-trivially on the set of qubits Q which is a subset of the total set of n qubits in the system. Then

$$\begin{aligned} E_\psi(P) &= \langle\psi|\left(\bigotimes_{p\in Q} \sigma_p\right)|\psi\rangle \\ &= \langle\psi|\left(\bigotimes_{p\in Q} \sigma_p\right)\left(\bigotimes_{q\notin Q} I_q\right)|\psi\rangle \\ &= \langle\psi|\left(\bigotimes_{p\in Q} B_{\sigma_p}^\dagger Z_p B_{\sigma_p}\right)\left(\bigotimes_{q\notin Q} I_q\right)|\psi\rangle \\ &= \langle\psi|\left(\bigotimes_{p\in Q} B_{\sigma_p}^\dagger\right)\left(\bigotimes_{p\in Q} Z_p\right)\left(\bigotimes_{q\notin Q} I_q\right)\left(\bigotimes_{p\in Q} B_{\sigma_p}\right)|\psi\rangle \end{aligned}$$

$$
\begin{aligned}
&= \langle\phi| \left(\bigotimes_{p\in Q} Z_p \right) \left(\bigotimes_{q\notin Q} I_q \right) |\phi\rangle \\
&= \langle\phi| \left(\bigotimes_{p\in Q} \sum_{x_p\in\{0_p,1_p\}} (-1)^{x_p} \left|x_p\right\rangle \left\langle x_p\right| \right) \left(\bigotimes_{q\notin Q} \sum_{y_q\in\{0_q,1_q\}} \left|y_q\right\rangle \left\langle y_q\right| \right) |\phi\rangle \\
&= \langle\phi| \left(\sum_{x\in\{0,1\}^n} (-1)^{\sum_{p\in Q} x_p} |x\rangle \langle x| \right) |\phi\rangle \\
&= \sum_{x\in\{0,1\}^n} (-1)^{\sum_{p\in Q} x_p} |\langle x|\phi\rangle|^2 \\
&= \sum_{x\in\{0,1\}^n} (-1)^{\sum_{p\in Q} x_p} P(|\phi\rangle \rightarrow |x\rangle), \qquad (3.48)
\end{aligned}
$$

where $|\phi\rangle = \left|\bigotimes_{p\in Q} B_{\sigma_p}\psi\right\rangle$. Finally, because the expectation value is linear

$$
E_\psi \left(\sum_m \lambda_m P_m \right) = \sum_m \lambda_m E_\psi(P_m), \qquad (3.49)
$$

one can estimate any observable that can be written as a linear combination of Puali-string terms.

3.4.3 Measurement

To estimate the probability $P(|\phi\rangle \rightarrow |x\rangle)$ from the previous section, one prepares the state $|\phi\rangle$ on a quantum computer and measures it, and then repeats this process (prepare and measure) several times. The probability $P(|\phi\rangle \rightarrow |x\rangle)$ is estimated to be the number of times that one measures the bit-string x divided by the total number of measurements that one makes; that is

$$
P(|\phi\rangle \rightarrow |x\rangle) \approx \sum_{m=1}^{M} \frac{x_m}{M}, \qquad (3.50)
$$

where

$$x_m = \begin{cases} 1 & \text{if the result of measurement is } x \\ 0 & \text{if the result of measurement is not } x. \end{cases} \tag{3.51}$$

By the law of large numbers [11], the approximation (3.52) approaches equality as M goes to infinity

$$P(|\phi\rangle \rightarrow |x\rangle) = \lim_{M\rightarrow\infty} \sum_{m=1}^{M} \frac{x_m}{M}. \tag{3.52}$$

As we obviously do not have infinite time nor infinite quantum computers to be run in parallel, we must truncate our number of measurement M to a finite, but sufficiently large number. More precisely, for precision ϵ, each expectation estimation subroutine within VQE requires $O(1/\epsilon^2)$ samples from circuits of depth $O(1)$ [85].

3.5 Transformations

While many-body nuclear physics operators are written in terms of fermionic operators, quantum computers work with Pauli operators. Thus, in order to simulate many-body nuclear physics on a quantum computer, we need a transformation between the two sets of operators. Several such transformations exist [74] (each with their own advantages and disadvantages) of which we list three here: Jordan-Wigner, Parity-Basis, and Bravyi-Kitiav.

Transformation	Basis	# of Operators	Locality
Jordan-Wigner	Occupation Number	$O(N)$	Local
Parity-Basis	Parity	$O(N)$	Local
Bravyi-Kitiav	Mixed	$O(\log(N))$	Non-local

Table 3.1: Comparison of basis, number of operators, and locality the of Jordan-Wigner, parity-basis, and Bravyi-Kitiav transformations.

Jordan-Wigner works in the occupation number representation which is naturally mapable to the computational basis set of a quantum computer. However, it requires long

strings of operators. Its main advantage for near-term devices is that it is local and hence implementable on quantum computers with linear qubit-connectivity. The parity-basis transformation is similar to Jordan-Wigner except that it works in the parity-basis. The Bravyi-Kitiav transformation is a mix between the Jordan-Wigner and the parity-basis transformations that allows for a shorter number of operators but which comes with the cost of being non-local and hence not suitable for quantum computers with limited qubit-connectivity. In this work, we use the Jordan-Wigner transformation because, although it requires more operators ($O(N)$) it uses the occupation number representation, which is naturally implementable on a quantum computer. Additionally, it is local, making it more easily implementable on devices with limited qubit connectivity, which describes most near-term devices. Here, qubit connectivity refers to which qubits are connected to which other qubits. Two qubits are connected if one can implement a two-qubit gate between them.

3.5.1 Jordan-Wigner Transformation

The Jordan-Wigner transformation was originally developed by Pascual Jordan and Eugene Wigner for one-dimensional lattice models [48]. The transformation is a mapping between fermionic and Pauli operators which stores information locally in the occupation number basis. It is given below as

$$a_p^\dagger = \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^-, \tag{3.53}$$

$$a_p = \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^+, \tag{3.54}$$

where

$$Q_p^\pm = \frac{X_p \pm iY_p}{2}, \tag{3.55}$$

and

$$\sigma_p = \left(\bigotimes_{n=1}^{p-1} I\right) \otimes \sigma \otimes \left(\bigotimes_{n=p+1}^{N} I\right), \tag{3.56}$$

where $\sigma = I, X, Y, Z$ is a Pauli operator and N is the number of qubits in the system. To gain some intuition for the mapping, note that the action on many-fermionic states (2.37) is preserved by the transformation

$$a_i^\dagger \left|n_1 \ldots n_n\right\rangle = \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^- \left|n_1 \ldots n_n\right\rangle = (-1)^{N_p}(1-n_p)\left|n_1 \ldots n_{p-1} 1 n_{p+1} \ldots n_n\right\rangle, \tag{3.57}$$

$$a_i \left|n_1 \ldots n_n\right\rangle = \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^+ \left|n_1 \ldots n_n\right\rangle = (-1)^{N_p} n_p \left|n_1 \ldots n_{p-1} 0 n_{p+1} \ldots n_n\right\rangle, \tag{3.58}$$

where $N_p = \sum_{m=1}^{p-1} n_m$, since

$$Z\left|n\right\rangle = (-1)^n \left|n\right\rangle, \tag{3.59}$$

$$Q^-\left|n\right\rangle = (1-n)\left|1\right\rangle, \tag{3.60}$$

$$Q^+\left|n\right\rangle = n\left|0\right\rangle, \tag{3.61}$$

where $n = 0, 1$. The mapping holds because it obeys the fermionic anti-commutation relations (2.39) as verified below: First, consider the case $p = q$. In this case, the anti-commutation relations are

$$\{a_p, a_p^\dagger\} = \left\{\left(\prod_{n=1}^{p-1} Z_n\right) Q_p^+, \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^-\right\} = \left(\prod_{n=1}^{p-1} \{Z_n, Z_n\}\right) \{Q_p^+, Q_p^-\} = I_p \tag{3.62}$$

$$\{a_p, a_p\} = \left\{\left(\prod_{n=1}^{p-1} Z_n\right) Q_p^+, \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^+\right\} = \left(\prod_{n=1}^{p-1} \{Z_n, Z_n\}\right) \{Q_p^+, Q_p^+\} = 0 \tag{3.63}$$

$$\{a_p^\dagger, a_p^\dagger\} = \left\{\left(\prod_{n=1}^{p-1} Z_n\right) Q_p^-, \left(\prod_{n=1}^{p-1} Z_n\right) Q_p^-\right\} = \left(\prod_{n=1}^{p-1} \{Z_n, Z_n\}\right) \{Q_p^-, Q_p^-\} = 0, \tag{3.64}$$

which follow from the fact that

$$\begin{aligned}\{Q_p^{\pm}, Q_p^{\mp}\} &= \frac{1}{4}\{X_p \pm iY_p, X_p \mp iY_p\} \\ &= \frac{1}{4}(\{X_p, X_p\} \mp i\{X_p, Y_p\} \pm i\{Y_p, X_p\} + \{Y_p, Y_p\}) = I_p, \end{aligned} \tag{3.65}$$

while

$$\begin{aligned}\{Q_p^{\pm}, Q_p^{\pm}\} &= \frac{1}{4}\{X_p \pm iY_p, X_p \pm iY_p\} \\ &= \frac{1}{4}(\{X_p, X_p\} \pm i\{X_p, Y_p\} \pm i\{Y_p, X_p\} - \{Y_p, Y_p\}) = 0. \end{aligned} \tag{3.66}$$

Second, consider the case $p \neq q$. Without loss of generality, we can set $p < q$. In this case, the anti-commutation relations are

$$\{a_p, a_q^{\dagger}\} = \left\{\left(\prod_{n=1}^{p-1} Z_n\right) Q_p^{+}, \left(\prod_{n=1}^{q-1} Z_n\right) Q_p^{-}\right\} \tag{3.67}$$

$$= \left(\prod_{n=1}^{p-1} \{Z_n, Z_n\}\right) \{Q_p^{+}, Z_p^{-}\} \left(\prod_{m=p+1}^{q-1} \{I_n, Z_m\}\right) \{I_q, Q_q^{-}\} = 0 \tag{3.68}$$

$$\{a_p, a_q\} = \left\{\left(\prod_{n=1}^{p-1} Z_n\right) Q_p^{+}, \left(\prod_{n=1}^{q-1} Z_n\right) Q_p^{+}\right\} \tag{3.69}$$

$$= \left(\prod_{n=1}^{p-1} \{Z_n, Z_n\}\right) \{Q_p^{+}, Z_p^{-}\} \left(\prod_{m=p+1}^{q-1} \{I_n, Z_m\}\right) \{I_q, Q_q^{+}\} = 0 \tag{3.70}$$

$$\{a_p^{\dagger}, a_q^{\dagger}\} = \left\{\left(\prod_{n=1}^{p-1} Z_n\right) Q_p^{-}, \left(\prod_{n=1}^{q-1} Z_n\right) Q_p^{-}\right\} \tag{3.71}$$

$$= \left(\prod_{n=1}^{p-1} \{Z_n, Z_n\}\right) \{Q_p^{-}, Z_p^{-}\} \left(\prod_{m=p+1}^{q-1} \{I_n, Z_m\}\right) \{I_q, Q_q^{-}\} = 0, \tag{3.72}$$

which follow from the fact that

$$\left\{Q_p^{\pm}, Z_p\right\} = \frac{1}{2}\{X \mp iY, Z\} = \frac{1}{2}\left(\{X, Z\} \mp i\{Y, Z\}\right) = 0 \tag{3.73}$$

3.5.2 Pair Jordan-Wigner Transformation

The Jordan-Wigner transformation is simplified when dealing with pair fermionic operators:

$$A_p^\dagger = Q_p^-, \tag{3.74}$$

$$A_p = Q_p^+, \tag{3.75}$$

$$N_p = I_p - Z_p. \tag{3.76}$$

Namely, the string of Z operators preceding the $Q^\pm$ operator is dropped. This is one of the main advantages of working with the nuclear pairing model. Because it can be written in terms of pair fermionic operators, its mapping to quantum operators is greatly simplified. This mapping holds because it obeys the pair fermionic commutation relations (5.6 - 5.8) as verified below:

$$\begin{aligned}
[A_p, A_q^\dagger] &= [Q_p^+, Q_q^-] \\
&= \frac{1}{4}[X_p + iY_p, X_q - iY_q] \\
&= \frac{1}{4}([X_p, X_q] - i[X_p, Y_q] + i[Y_p, X_q] + [Y_p, Y_q]) \\
&= \delta_{pq} Z_p \\
&= \delta_{pq}(I_p - N_p),
\end{aligned} \tag{3.77}$$

and

$$\begin{aligned}
[N_p, A_q^\dagger] &= [I_p - Z_p, Q_q^-] \\
&= \frac{1}{2}[I_p - Z_p, X_q - iY_q] \\
&= \frac{1}{2}([I_p, X_q] - i[I_p, Y_q] - [Z_p, X_q] + i[Z_p, Y_q])
\end{aligned}$$

$$= \delta_{pq}\left(\frac{X_p - iY_p}{2}\right)$$
$$= \delta_{pq}A_p^\dagger, \tag{3.78}$$

and

$$\begin{aligned}
[N_p, A_q] &= [I_p - Z_p, Q_q^+] \\
&= \frac{1}{2}[I_p - Z_p, X_q + iY_q] \\
&= \frac{1}{2}([I_p, X_q] + i[I_p, Y_q] - [Z_p, X_q] - i[Z_p, Y_q]) \\
&= \delta_{pq}\left(\frac{X_p + iY_p}{2}\right) \\
&= \delta_{pq}A_p. \qquad (3.79)
\end{aligned}$$

CHAPTER 4

LIPKIN MODEL

4.1 Introduction

The Lipkin model is an exactly solvable, many-body toy-model, first introduced in 1965 by Lipkin, Meschkov, and Glick [56]. It is often used to test the validity of nuclear many-body methods. The version we consider here describes pairing interactions between two levels (with the same j-value that straddle the Fermi level). The model consists of N nucleons, distributed over two, Ω-degenerate levels which are indexed by $\sigma = \pm 1$; here $\Omega = 2j + 1$. The model is depicted schematically in Figure 4.1 below. The solid lines represent the available energy levels while the dashed line represents the Fermi-level. The model is described by the following Hamiltonian

$$H = \frac{1}{2}\epsilon \sum_{n\sigma} \sigma a^{\dagger}_{n\sigma} a_{n\sigma} - \frac{1}{2} V \sum_{nm\sigma} a^{\dagger}_{n\sigma} a^{\dagger}_{m\sigma} a_{m\bar{\sigma}} a_{n\bar{\sigma}}, \tag{4.1}$$

where $n, m = 1, 2, ..., \Omega$ and $\sigma = \pm 1$ (with $\bar{\sigma} = -\sigma$). The single-particle energy e is the amount of energy required to move a nucleon between the lower-level ($\sigma = -1$) which has energy $-\epsilon/2$ to the upper-level ($\sigma = +1$) which has energy $\epsilon/2$. Additionally, the interaction strength V is the energy required to move a pair of nucleons between the lower and upper levels. Note that in this model, nucleons must move between levels in pairs, either two in the lower level moving together to the upper or vice verse; this is why it is

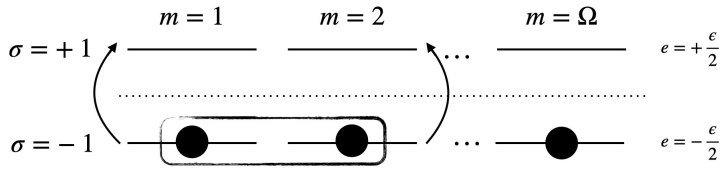


Figure 4.1: Schematic of Lipkin Model.

described as a pairing model.

4.2 Classical Solutions

4.2.1 Full Configuration Interaction

The Lipkin model can be exactly solved via the full configuration interaction (FCI) method described in subsection (2.4.1). The FCI basis consists of the Slater determinants

$$|n\rangle = |n_1 \cdots n_{2\Omega}\rangle = (a_1^\dagger)^{n_1} ... (a_{2\Omega}^\dagger)^{n_{2\Omega}} |0\rangle \,, \tag{4.2}$$

where $n_k = 0, 1$ is the occupation number of the state with index $m = \lfloor k/2 \rfloor$ and $\sigma = 2[k \pmod 2] - 1$, for $k = 1, 2, ..., 2\Omega$. That is, even (odd) values of k index the upper (lower) level, respectively. For N nucleons, the FCI basis consists of all states $|n\rangle$ with a Hamming weight of N; that is, $n_k = 1$ for N values of k. The size of the basis is thus $\binom{2\Omega}{N}$. The diagonal Hamiltonian matrix elements are given by

$$\langle m_1 \cdots m_{2\Omega}|H|n_1 \cdots n_{2\Omega}\rangle = \frac{\epsilon}{2} \left(N_+ - N_-\right), \tag{4.3}$$

where $n_k = m_k$ for all k. Here

$$N_+ = \sum_{k=0}^{\Omega-1} n_{2k+1}, \tag{4.4}$$

$$N_- = \sum_{k=0}^{\Omega-1} n_{2k}, \tag{4.5}$$

are the number of nucleons in the upper and lower levels, respectively. The off-diagonal matrix elements are given by

$$\langle m_1 \cdots m_{2\Omega}|H|n_1 \cdots n_{2\Omega}\rangle = -\frac{V}{2}, \tag{4.6}$$

where $n_{2k} n_{2j} n_{2k+1} n_{2j+1} = x x \bar{x} \bar{x}$ and $m_{2k} m_{2j} m_{2k+1} m_{2j+1} = \bar{x} \bar{x} x x$ for exactly one pair (k, j) from $k \neq j = 0, ..., \Omega - 1$, and $n_l = m_l$ for all $l \neq 2k, 2k+1, 2j, 2j+1$. Here, $x = 0, 1$

with $\bar{0} = 1$ and $\bar{1} = 0$. The eigenvalues and eigenvectors are then found through direct diagonalization of the Hamiltonian matrix H which has elements

$$H_{n_1 \ldots n_{2\Omega}, n_1 \ldots n_{2\Omega}} = \langle m_1 \cdots m_{2\Omega} | H | n_1 \cdots n_{2\Omega} \rangle \,. \tag{4.7}$$

The eigenvalue energies are computed for the case $\Omega = N = 4$ against various values of the interaction strength V. They are depicted as lines in Figure 4.2. This was the largest value of Ω that could be solved in a reasonable amount of time on my laptop, underscoring the exponential time-scaling of the FCI method.

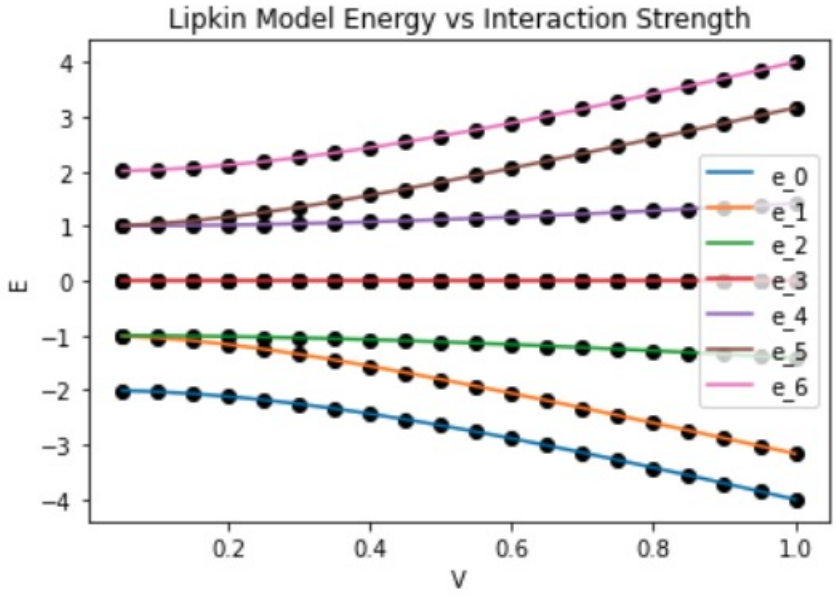


Figure 4.2: The energy eigenvalues (E) of the Lipkin model, are plotted for various interaction strengths (V). The level degeneracy Ω and particle number N are both four while the single-particle energy ϵ is one. The different energies e_k for $k = 0, 1, ..., 6$ are depicted by different colors, labeled in the plot itself. The solid lines are the results of the FCI method while the dots are the results of the symmetry method.

4.2.2 Symmetry Method

Following the procedure of the symmetry method laid out in subsection (2.4.2), we start by identifying the symmetries of the Lipkin Hamiltonian. The first symmetry is the particle-number. The particle number operator

$$N = \sum_{n\sigma} a^{\dagger}_{n\sigma} a_{n\sigma}, \tag{4.8}$$

commutes with the Lipkin Hamiltonian. This can be seen by examining the Hamiltonian (4.1) and noticing that the one-body part simply counts particles while the two-body term moves particles in pairs. Thus, the Hamiltonian conserves particle number. To find more symmetries, we rewrite the Lipkin Hamiltonian in terms of SU(2) operators

$$H = \epsilon J_z + \frac{1}{2} V (J_+^2 + J_-^2), \tag{4.9}$$

via the mapping

$$J_z = \sum_n j_z^{(n)}, \tag{4.10}$$

$$J_\pm = \sum_n j_\pm^{(n)}, \tag{4.11}$$

where

$$j_z^{(n)} = \frac{1}{2} \sum_\sigma \sigma a^{\dagger}_{n\sigma} a_{n\sigma}, \tag{4.12}$$

$$j_\pm^{(n)} = a^{\dagger}_{n\pm} a_{n\mp}. \tag{4.13}$$

These operators obey the SU(2) commutation relations

$$[J_+, J_-] = 2J_z, \tag{4.14}$$

$$[J_z, J_\pm] = \pm J_\pm, \tag{4.15}$$

as justified in Appendix A. Here, the ladder operators are defined as $J_\pm = J_x \pm iJ_y$. With this rewriting, we can see that the total spin operator J^2, which is defined as

$$J^2 = J_x^2 + J_y^2 + J_z^2 = \frac{1}{2}\{J_+, J_-\} + J_z^2, \tag{4.16}$$

commutes with the Hamiltonian since the Hamiltonian is written explicitly in terms of SU(2) operators and J^2 is the center of SU(2), meaning that it commutes with all the group's elements. Finally, we note that the signature operator

$$R = e^{i\pi J_z}, \tag{4.17}$$

commutes with the Hamiltonian, which can be explained as follows: Writing J_z as

$$J_z = \frac{1}{2}(N_+ - N_-), \tag{4.18}$$

where $N_\pm = \sum_{n\pm} a_{n\pm}^\dagger a_{n\pm}$, allows us to see that it measures half the difference between the number of particles in the upper and lower levels. Thus, the possible eigenvalues r of the signature operator are

$$r = \begin{cases} +1, & j_z = 2n \\ +i, & j_z = 2n + \frac{1}{2} \\ -1, & j_z = 2n + 1 \\ -i, & j_z = 2n + \frac{3}{2} \end{cases} \tag{4.19}$$

for $n \in \mathbb{Z}$. Note that r is real or imaginary if the number of particles N is even or odd, respectively. Since, as discussed above, the Lipkin Hamiltonian conserves N, r cannot jump between being real and imaginary. Additionally, because particles must be moved in

pairs, and J_z measures half the difference between particles in the upper and lower levels, j_z can only change by as

$$\begin{aligned} j_z &\rightarrow \frac{1}{2}[(N_+ \pm 2n) - (N_- \mp 2n)] \\ &= J_z \pm 2n. \end{aligned} \tag{4.20}$$

We have determined that the symmetry operators N, J, and R commute with the Hamiltonian. Let their eigenvalues be n, j, and r, respectively. These become our new quantum numbers. Starting with the particle number operator N, because H cannot mix states from one particle number to another ($\langle N|H|N'\rangle = 0$ if $N \neq N'$) the Hamiltonian matrix is block diagonal with N blocks. Each block corresponds to a different particle number ($n = 0, 1, ..., 2\Omega$) with size $n^2/4+n$. This is a direct result of the particle number symmetry. We now move on to the total-spin operator. Because the Hamiltonian cannot mix states with different J, we have that $\langle J|H|J'\rangle = 0$ for $J \neq J'$. Now, for a given N, we label our basis states $|j j_z\rangle$ where $j_z = -j, -j+1, ..., j-1, j$. The non-zero Hamiltonian matrix elements are

$$\langle JJ_z| H \left|J'J'_z\right\rangle = \delta_{JJ'}\epsilon j_z, \tag{4.21}$$

$$\begin{aligned} \langle JJ_z + 2| H |JJ_z\rangle &= -\frac{V}{2}\sqrt{[j(j+1) - j_z(j_z+1)]} \\ &\quad \times \sqrt{[j(j+1) - (j_z+1)(j_z+2)]}, \end{aligned} \tag{4.22}$$

$$\begin{aligned} \langle JJ_z - 2| H |JJ_z\rangle &= -\frac{V}{2}\sqrt{[j(j+1) - j_z(j_z-1)]} \\ &\quad \times \sqrt{[j(j+1) - (j_z-1)(j_z-2)]}, \end{aligned} \tag{4.23}$$

since the operators that make up the Hamiltonian act on the basis states as follows:

$$J_z |JJ_z\rangle = j_z |JJ_z\rangle , \tag{4.24}$$

$$J_{\pm} |JJ_z\rangle = \sqrt{j(j+1) - j_z(j_z \pm 1)}\, |JJ_z \pm 1\rangle\,, \tag{4.25}$$

Note that the maximum possible value of j_z is $N/2$ which would correspond to the state where all N particles are spin up:

$$j_z |\uparrow \cdots \uparrow\rangle = \frac{1}{2} \sum_{m\sigma} N_{m\sigma} |\uparrow \cdots \uparrow\rangle = \frac{N}{2}. \tag{4.26}$$

Therefore, the maximum value of j is also $N/2$ and thus its possible values are $j = N/2, N/2-1, ..., 1$ if N is even and $j = N/2, N/2-1, ..., 1/2$ if N is odd. For each j, the possible values of j_z are $j_z = j, j-1, ..., -j$. Thus, there are $2j+1$ possible values of j_z for each j. This implies that each n-block of the Hamiltonian matrix is itself a block diagonal matrix consisting of $\lfloor N/2 \rfloor$ blocks. Each block corresponds with a total spin value ($j = N/2, N/2-1, ...$) and has length $2j+1$. This is the direct result of the total spin symmetry. Finally, we move to the signature R. Each j-block is again, itself a block diagonal with two blocks ($r = \pm 1$ if N is even or $r = \pm i$ if N is odd) which have size j and $j+1$, respectively. The energies are computed by direct diagonalization of the Hamiltonian matrix we've been describing above.

One can see in Figure 4.2 that the energies computed via the symmetry method (depicted as block dots) exactly match the results of the FCI method. This is a demonstration that the symmetry method is indeed a valid, exact solution. The symmetry method is also used to solve the Lipkin Hamiltonian for the case of $\Omega = N = 10$. With the FCI method, this would involve diagonalizing a size $\binom{20}{10} \sim 10^5$, square matrix. But, with the symmetry method, one need only diagonalize several smaller square matrices, the largest of which is size 30, which is $j(j+1)$ for $j = N/2$. The eigenvalues of the Lipkin Hamiltonian for single-particle energy $\epsilon = 1$ are plotted against various pairing strengths V in Figure 4.3.

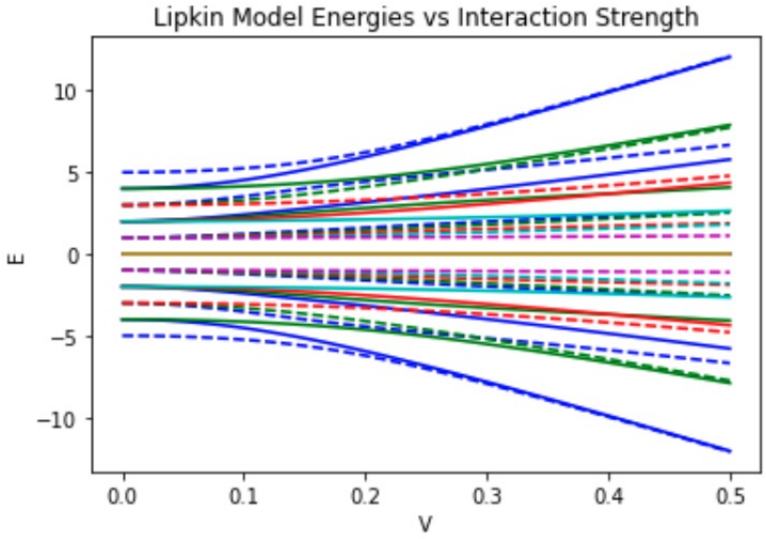


Figure 4.3: The energy eigenvalues (E) of the Lipkin model, computed via the symmetry method, are plotted for various interaction strengths (V). The level degeneracy Ω and particle number N are both ten while the single-particle energy ϵ is one. The solid and dashed lines correspond to signature numbers $r = +1$ and $r = -1$, respectively. The colors yellow, magenta, cyan, red, green, and blue correspond to $j = 0, 1, ..., 5$, respectively.

Each color represents a different value of j. For example, the eigenvalue energies plotted in blue were the result of diagonalizing the $j = 5$ block. The lines are solid and dashed to correspond to the signature $r = +1$ and $r = -1$, respectively. We note that for $V = 0$, there are $11 = N + 1$ values that the energies can take, corresponding to the fact that the Lipkin model with no interaction strength is simply ϵJ_z; thus the Hamiltonian simply counts half the difference between the number of particles in the two levels, a number which has 11 possible values: $j_z = 0, \pm1, \pm2, \pm3, \pm4, \pm5$. However, as the pairing strength is turned on, the energies start to bend and split. We notice that as V increase, the energies for $r = +1$ and $r = -1$ start to pair up and equal one another, their states becoming degenerate.

4.2.3 Hartree-Fock Method

As mentioned in subsection 2.4.3, the ansatz we use for the variational method is the particle-number conserving product state which, for the Lipkin model, is labeled as

$$|\tau\rangle = \exp\left(\sum_{n} \tau_{n_+n_-} a^\dagger_{n+} a_{n-}\right) \prod_{k=1}^{N} a^\dagger_k \,|0\rangle\,, \tag{4.27}$$

where $\tau_{n_+n_-}$ is a variational parameter. This is motivated by Thouless's theorem which states that such an ansatz can rotate any Slater determinant into any other. Since we are only considering single Slater determinants, this is exactly what we desire. Here, we consider the half-filled case $N = \Omega$. Thus, we can start the ansatz in the state

$$\prod_{k=1}^{N} a^\dagger_k \,|0\rangle = |J - J\rangle\,. \tag{4.28}$$

However, because the Lipkin Hamiltonian treats all $(n+, n-)$ pairs equivalently (the two-body coefficient V is independent of n and σ) we can set

$$\tau_{n_+n_-} = \tau, \tag{4.29}$$

a new variational parameter, for all n. With this, the normalized ansatz becomes

$$|\tau\rangle = (1 + |\tau|^2)^J e^{\tau J_+} \,|0\rangle\,, \tag{4.30}$$

where the normalization is determined in Appendix B. We now calculate the expectation value of the Hamiltonian in the ansatz. Using Appendix C, we derive

$$\begin{aligned} E(\tau) = \langle\tau|H|\tau\rangle &= \frac{\Omega}{2}\left[\epsilon\frac{|\tau|^2 - 1}{|\tau|^2 + 1} - V(\Omega - 1)\frac{\tau^2 + \bar{\tau}^2}{\left(|\tau|^2 + 1\right)^2}\right], \\ &= -\frac{1}{2}\epsilon\Omega\left(\cos\theta + \frac{\chi}{2}\sin^2\theta\cos 2\phi\right), \end{aligned} \tag{4.31}$$

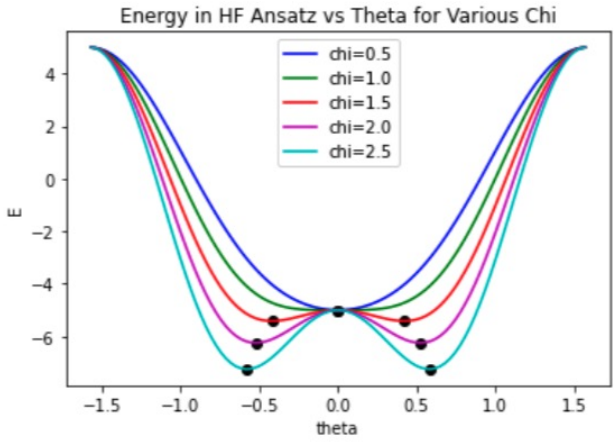


Figure 4.4: The expectation value of the Lipkin Hamiltonian in the SU(2) coherent state ansatz (4.30) is plotted vs theta for various values of χ which are distinguished by different colors, labeled on the plot itself. The black dots represent the minimum energy (4.38).

where we've defined the coupling strength χ to be

$$\chi = \frac{V}{\epsilon}(\Omega - 1). \tag{4.32}$$

This energy profile $E(\tau) = \langle\tau|H|\tau\rangle$ is plotted for various values of χ in Figure 4.4. The plot shows the symmetry breaking that occurs when χ becomes greater than 1. For $\chi \leq 1$, there is only one θ_{min} (namely $\theta_{\text{min}} = 0$). However, for $\chi > 1$, there exists two values of θ_{min} which give the same, correct ground-state energy, and are symmetric about $\theta = 0$. To minimize $E(\tau)$, we first set its derivatives to zero, resulting in the following expressions:

$$0 = \frac{\partial E}{\partial \theta} = \frac{1}{2}\epsilon\Omega \sin\theta(1 - \chi\cos\theta\cos 2\phi), \tag{4.33}$$

$$0 = \frac{\partial E}{\partial \phi} = \frac{1}{2}\epsilon\Omega\chi \sin^2\theta \sin 2\phi, \tag{4.34}$$

the first of which implies either $\theta_{\text{min}} = 0, \pi$ or $\cos\theta_{\text{min}}\cos 2\phi_{\text{min}} = 1/\chi$ and the second of which implies either $\theta_{\text{min}} = 0, \pi$ or $\phi_{\text{min}} = 0, \pi/2$. Second, we demand that its second derivatives are positive

$$0 < \frac{\partial^2 E}{\partial \theta^2} = \frac{1}{2}\epsilon\Omega(\cos\theta - \chi\cos 2\theta\cos 2\phi), \tag{4.35}$$

$$0 < \frac{\partial^2 E}{\partial \phi^2} = \epsilon \Omega \chi \sin^2\theta \cos 2\phi, \tag{4.36}$$

which implies (assuming, without loss of generality, that $\chi > 0$) $-\pi/3 < \theta_{\min} < \pi/3$ and $\phi_{\min} < \pi/4$. And third, we set the cross-derivative equal to zero

$$0 = \frac{\partial^2 E}{\partial \theta \partial \phi} = \epsilon \Omega \chi \cos 2\theta \sin 2\phi, \tag{4.37}$$

which implies $\theta_{\min} = 0, \pi$. Combining all these conditions implies

$$\theta_{\min} = \begin{cases} 0, & \text{if } 0 < \chi \leq 1 \\ \cos^{-1}\left(\frac{1}{\chi}\right), & \text{if } 1 < \chi \end{cases}, \tag{4.38}$$

$$\phi_{\min} = 0. \tag{4.39}$$

The minimum variational parameter is thus, $\tau_{\min} = \tau(\theta_{\min}, \phi_{\min})$ where $\tau(\theta, \phi) = \tan(\theta/2)e^{-i\phi}$, as defined in Appendix C. Finally, we plug these minimum parameters (4.38 and 4.39) back into $E(\tau)$ (4.31) to find the Hartree-Fock ground state energy to be

$$E_0 = \begin{cases} -\frac{\epsilon}{2}\Omega, & \text{if } 0 < \chi < 1 \\ -\frac{\epsilon}{4}\Omega\left(\chi + \frac{1}{\chi}\right), & \text{if } 1 < \chi. \end{cases} \tag{4.40}$$

The Hartree-Fock method is bench-marked against the exact answer in Figure 4.5. The Hartree-Fock calculated ground-state energy (4.40) is plotted against various values of the coupling strength χ. Alongside it we plot the exact ground state energy computed using the symmetry method. We more precisely inspect the performance of the Hartree-Fock method by plotting the relative error between the Hartree-Fock and exact energies in Figure 4.6. We note that the two methods start in exact agreement at $\chi = 0$. The magnitude of the relative error increases until just after 1. It then decreases as χ increases past 1 and

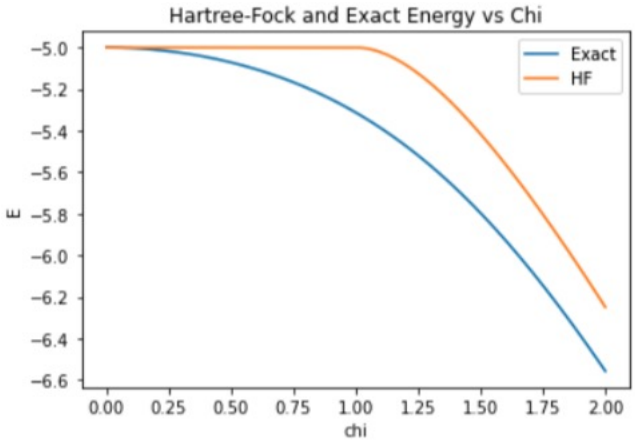


Figure 4.5: The Hartree-Fock and exact energies of the Lipkin are plotted against various values of χ.

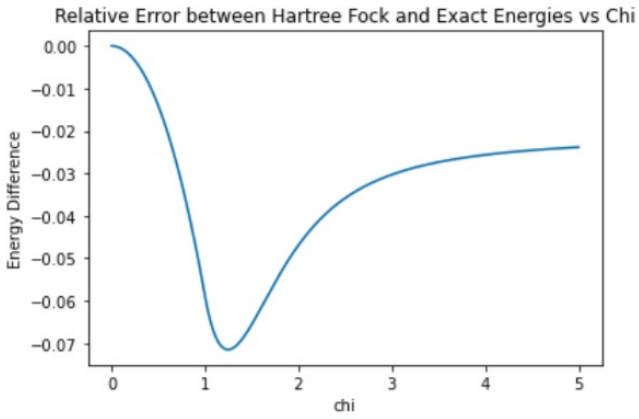


Figure 4.6: The relative difference between the Hartree-Fock and exact energies is plotted against various values of χ.

seems to asymptotically approach a small error. Thus, one could say that the Hartree-Fock method is a good approximation for either very small or very large χ.

4.3 Quantum Solutions

To solve the Lipkin model with a quantum computer, the first step is to map the system to a set of qubits. We'll restrict ourselves here to the half-filled case where the number of particles N equals the degeneracy of the states Ω. One could assign each possible state

(n, σ) to a qubit such that the qubit being in the state $|1\rangle$ or $|0\rangle$ would imply that the state (n, σ) is occupied or unoccupied, respectively. This mapping scheme (which we'll call occupation mapping) requires 2Ω qubits. Additionally, any ansatz that would restrict the minimization search to the correct subspace of constant Hamming weight N (since the number of particles N is conserved) would necessitate the use of at least four-qubit gates. This is because moving a pair of particles in this scheme would require two annihilation operators on the states from which the pair particles move and two creation operators on the states to which the pair of particles move. That is, it takes a four-qubit gate to change between the states $|1100\rangle$ and $|0011\rangle$, for example. And, as discussed in the chapter of quantum computing, it is only known how to efficiently decompose up to two-qubit gates. Thus the involvement of four qubit gates would necessitate a longer depth circuit than one involving only two and one qubit gates, creating more noise and less accurate results.

However, because there are only two energy levels in the Lipkin model, another natural mapping is possible. In this mapping scheme (which we'll call level mapping) each doublet $((n, +1), (n, -1))$ would be assigned a qubit such that the qubit being in the state $|0\rangle$ or $|1\rangle$ would imply that the particle is in the $(n, +1)$ or $(n, -1)$ state, respectively. Note that these are the only two possible configurations of the doublet as we are restricting ourselves to the half-filled case and the Lipkin Hamiltonian only moves particles between energy levels, not degenerate states. Thus the level mapping only requires Ω qubits which is half that of the occupation mapping. Additionally, any ansatz that would restrict the minimization search to the correct subspace of constant Hamming weight N requires at most, only two-qubit gates. This is because moving a pair of particles in this scheme only changes the state of two doublets (and therefore qubits). That is, it only takes a two-qubit gate to change

between the states $|00\rangle$ and $|11\rangle$, for example. As an efficient decomposition two-qubit gates is known, the ansatz for this mapping would be shorter (and thus less noisy) than that of the previous mapping.

One could imagine a third mapping scheme which would require even less qubits in which each of the possible states in the spin basis $|JJ_z\rangle$ is mapped to a single qubit. In this spin mapping, there are only $2J + 1$ possible states (since $J_z = -J, -J + 1, ..., J - 1, J$) for each value of J. And, since the Hamiltonian is block diagonal (with a different block for each J) the eigenvalues of the Hamiltonian are simply the eigenvalues of each block, which may be calculated separately. Since the maximum value of J is $J_{\max} = N/2$, the largest number of qubits would be $2J_{\max} + 1 = N + 1$. However, $\lfloor N/2 \rfloor$ different circuit would need to be used for minimization for all possible values of J, to explore the entire Hilbert space. (The minimum of the set of minimum energies that each circuit finds would be the ground state energy of the entire system.) This increases, linearly, the amount of time required to find the ground state energy.

After reviewing the three possible mappings, it is our view that the level mapping [17] is the best suited for NISQ era devices given its low qubit count and ability to search the entire relevant Hilbert space with one circuit (which reduces time to solution) and the fact that at most, only two-qubit gates are required of the ansatz, leading to shorter depth (and thus less noisy) circuits. With this mapping, the Hamiltonian takes the form utilized in the symmetry method (subsection 4.2.2) which was given by equation (4.9) as

$$H = \epsilon J_z + \frac{1}{2} V (J_+^2 + J_-^2). \tag{4.41}$$

Plugging the mapping from the total J operators to individual j operators (equations 4.12

and 4.13) yields

$$H = \epsilon \sum_n j_z^{(n)} + \frac{1}{2} V \left[\left(\sum_n j_+^{(n)} \right)^2 + \left(\sum_n j_-^{(n)} \right)^2 \right] \tag{4.42}$$

$$= \epsilon \sum_n j_z^{(n)} + \frac{1}{2} V \sum_{n,m} \left(j_+^{(n)} j_+^{(m)} + j_-^{(n)} j_-^{(m)} \right) \tag{4.43}$$

$$= \epsilon \sum_n j_z^{(n)} + 2V \sum_{n<m} \left(j_x^{(n)} j_x^{(m)} - j_y^{(n)} j_y^{(m)} \right), \tag{4.44}$$

where we've used the definitions

$$j_\pm^{(n)} = j_x^{(n)} \pm i j_y^{(n)}. \tag{4.45}$$

To convert to Pauli matrices, we'll make the transformations

$$j_x^{(n)} \to X_n/2, \tag{4.46}$$

$$j_y^{(n)} \to Y_n/2, \tag{4.47}$$

$$j_z^{(n)} \to Z_n/2, \tag{4.48}$$

which preserves the SU(2) commutation relations (4.14 and 4.15) and thus is allowable. This transforms our Hamiltonian into

$$H = \frac{1}{2} \epsilon \sum_{k=1}^{n} Z_k + \frac{1}{2} V \sum_{i \neq j=1}^{N} (X_k X_j - Y_k Y_j). \tag{4.49}$$

With this form, we can clearly see that the first (one-body) term in the Hamiltonian returns the energy $-\epsilon/2$ or $+\epsilon/2$ if the qubit representing the particle of a doublet is in the ground ($|1\rangle$) or excited ($|0\rangle$) state, respectively. The action of the second (two-body) term in the Hamiltonian can be determined by noting that

$$\frac{1}{2}(XX - YY) \, |00\rangle = |11\rangle \,, \tag{4.50}$$

$$\frac{1}{2}(XX - YY)\,|01\rangle = 0, \tag{4.51}$$

$$\frac{1}{2}(XX - YY)\,|10\rangle = 0, \tag{4.52}$$

$$\frac{1}{2}(XX - YY)\,|11\rangle = |00\rangle\,. \tag{4.53}$$

That is, the two-body term moves a pair of particles between the ground states $|00\rangle$ and the excited states $|11\rangle$ of their respective doublets.

To construct an efficient ansatz, we must determine the subspace within which the Hamiltonian lives. To begin, note that particles are only ever moved between energy levels in pairs. This implies that all possible states have a Hamming weight of constant parity (odd or even); this is the same as the signature r being conserved. Further, note that the Hamiltonian's coefficients (ϵ and V) are state independent (do not depend on the indices n or m) as the states labeled by these indices are degenerate and thus have the same energy level. Thus, the Hamiltonian treats all states with the same number of excited particles (Hamming weight of the state) as the same. Therefore, the following ansatz forms exactly cover the subspace within which the N-degenerate Hamiltonian explores:

$$|\psi_{\text{even}}\rangle = \sum_{k=0}^{\lfloor n/2 \rfloor} c_{2k} \left|D_{2k}^{n}\right\rangle, \tag{4.54}$$

$$|\psi_{\text{odd}}\rangle = \sum_{k=0}^{\lfloor n/2 \rfloor} c_{2k+1} \left|D_{2k+1}^{n}\right\rangle. \tag{4.55}$$

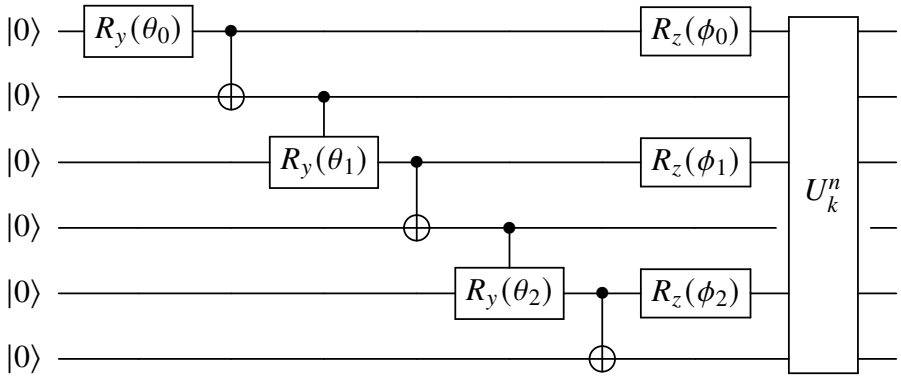
Here $\left|D_{k}^{n}\right\rangle$ represents a Dicke state which is defined as equal superposition of all n-qubit states with Hamming weight k. That is

$$\left|D_{k}^{n}\right\rangle = \frac{1}{\sqrt{\binom{n}{k}}} \sum_{x \in h_{k}^{n}} |x\rangle\,, \tag{4.56}$$

where $h_k^n = \{|x\rangle \mid \mathrm{l}(x) = n,\ \mathrm{wt}(x) = k\}$. One way to prepare these ansatzes is to do so exactly as it is known how to deterministically prepare Dicke states with linear depth [9]. The reference provides an algorithm for preparing a set of gates U_k^n that prepares a Dicke state from a product state of Hamming weight k; that is

$$U_k^n |1\rangle^{\otimes k} |0\rangle^{\otimes n-k} = \left|D_k^n\right\rangle . \tag{4.57}$$

It then describes how to one can create an arbitrary superposition of Dicke states, which we modify here to restrict ourselves to a Hamming weight of constant parity. The circuit to construct such a state (for the $k = 6$ case, as an example) is given below



(4.58)

The R_y gates and CNOT gates prepare an arbitrary real superposition of product states with even Hamming weight k; then the R_z gates add arbitrary phases to each of the states

$$\begin{aligned} |000000\rangle \rightarrow\ & \cos(\theta_0/2)\,|000000\rangle \\ &+ \sin(\theta_0/2)\cos(\theta_1/2)e^{i\theta_0}\,|110000\rangle \\ &+ \sin(\theta_0/2)\sin(\theta_1/2)\cos(\theta_2/2)e^{i(\theta_0+\theta_1)}\,|111100\rangle \\ &+ \sin(\theta_0/2)\sin(\theta_1/2)\sin\theta_2/2)e^{i(\theta_0+\theta_1+\theta_2)}\,|111111\rangle . \end{aligned} \tag{4.59}$$

Finally, U_k^n converts each product state to its corresponding Dicke state. Thus, all together the circuit acts as

$$\begin{aligned}|000000\rangle \rightarrow\ & \cos(\theta_0/2)\left|D_0^6\right\rangle \\ &+\ \sin(\theta_0/2)\cos(\theta_1/2)e^{i\theta_0}\left|D_2^6\right\rangle \\ &+\ \sin(\theta_0/2)\sin(\theta_1/2)\cos(\theta_2/2)e^{i(\theta_0+\theta_1)}\left|D_4^6\right\rangle \\ &+\ \sin(\theta_0/2)\sin(\theta_1/2)\sin\theta_2/2)e^{i(\theta_0+\theta_1+\theta_2)}\left|D_6^6\right\rangle . \end{aligned} \tag{4.60}$$

The circuit (4.58) can be extended naturally for any even value of k. For odd values of k, one need simply add a single-qubit to the top of the circuit for $k-1$ and apply the X gate to it. Although this ansatz has linear depth, the circuit for U_k^n involves several double-controlled gates which involve the usage of several CNOT gates to decompose. As the CNOT gate is often the noisiest gate in NISQ era quantum computers, it is best to minimize their use.

4.4 Results

In this section, we test out ansatz 4.58 for the Lipkin model with parameters. $\Omega = 4$, $e = 1$ and $v = 1$. One can see if Figure 4.7 that running VQE with our ansatz matches the exact energy for the most part, and always performs better than Hartree-Fock. Because the simulations of VQE were noiseless, we hypothesize that the slight variations in some of the VQE dots (red) off of the exact energy line (blue) could be due to the minimization algorithm failing to converge properly. Finding a set of initial parameters that would initialize us to a state with a large overlap with the ground state would be beneficial.

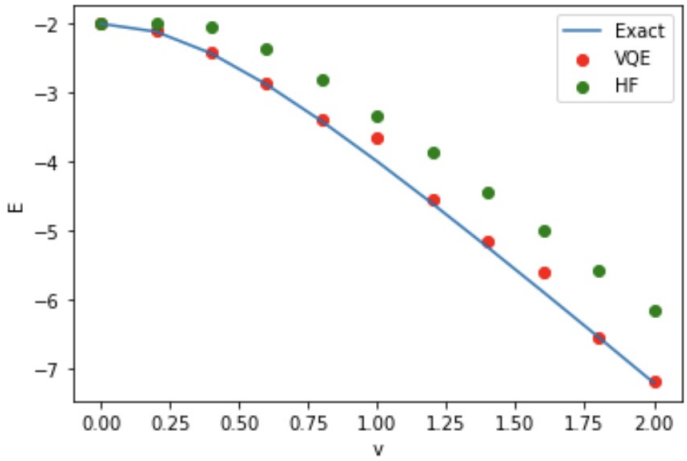


Figure 4.7: Comparison of energies for the Lipkin model calculated through direct diagonalization (Exact), Hartree-Fock (HF), and the variational quantum eigensolver (VQE).

4.5 Conclusion

In this section, we introduced the Lipkin model which serves as a toy model in nuclear physics with which to benchmark new techniques. We first solve the problem through various classical avenues including the full configuration interaction, symmetry, and Hartree-Fock methods. We then discussed the different ways to map the problem from its fermionic space to the spin space with which which quantum computers deal. We gave a novel way to construct one form of the ansatz for the model, implemented it as the ansatz for VQE and compared the quantum results to the classical results. This section served as a first example of how quantum computers can be used to solve a toy nuclear pairing model.

CHAPTER 5

PAIRING MODEL

5.1 Introduction

There exists an intriguing phenomenon in physics called pairing in which fermions "pair up", meaning that they move close together in space and tend to move energy levels together. The phenomenon of nuclear pairing can be understood from a simple symmetry argument. Since nucleons are fermions, the overall wavefunction of a pair nucleons must be anti-symmetric. Any such wavefunction can be written as the product of three separable wavefunctions: a spatial function, a spin function, and an isospin function. Pairing occurs between two fermions of opposite spin, which necessitates that the spin ($S = 0$) function of their combined system be anti-symmetric. Additionally, pairing occurs most strongly between two nucleons of the same type (proton-proton or neutron-neutron) which is described by an isospin of $T = 1$, necessitating that the isospin function of their combined system be symmetric. The spin function being anti-symmetric and the isospin function being symmetric implies that the spatial function must be symmetric (to preserve the overall anti-symmetry of the pair). This occurs when the angular momentum component of the spatial wavefunction is zero, $l = 0$, resulting in a wavefunction whose density has a strong peak near $r = 0$, where r is the separation between the two nucleons. Thus, the two nucleons tend to stick close together and can be approximated as a pair that moves together. Pairing occurs most strongly when $J = 0$ because $J = S + L$ and we determined above that $L = S = 0$. Because of this, a large energy gap is produced between the $J = 0$ and $J > 0$ states by the nuclear force between identical nucleons. Therefore, one may approximate such a system via the so-called pairing interaction which only acts

on the paired $J = 0$ state [22]. As we are approximating the residual interaction between identical nucleons, the pairing approximation is only suitable for semi-magic nuclei with valence nucleons of a single type [47]. Indeed, this paring approximation has been used to approximate such systems in heavy nuclei [27]. It will be seen that models that include only pairing interactions in their two-body forces are computationally simpler to solve and yet, as just mentioned, still applicable to real-world systems. Thus, such models serve as an area of interest for applications of NISQ era quantum computing. The pairing interaction was first introduced by Racah for atomic physics [65]. The algebra of identical nucleon pairs (isospin $t = 0$) is found to be isomorphic to SU(2) and is thus called quasi-spin. This is advantages for quantum computers whose qubits live in SU(2). One of the earliest applications of pairing was by Bardeen, Cooper, and Schrieffer in their famous model (BCS) of superfluidity in condensed matter in 1957 [7]. The idea was adapted to pairing in nuclei by Bohr, Mottelson, and Pines in 1958 [12].

In 1963, R.W. Richardson proposed a model consisting of fermions occupying non-degenerate energy levels which interact solely through the pairing force ([1] and [66]). We refer here to his model as the pairing model. It consists of P, non-degenerate energy levels, occupied by N pairs of fermions. Each pair consists of two fermions of opposite spin, occupying the same energy level. It's Hamiltonian is given by

$$H_p = \sum_{p\sigma} d_p a^\dagger_{p\sigma} a_{p\sigma} + \sum_{pq} g^p_q a^\dagger_{p+} a^\dagger_{p-} a_{q-} a_{q+}. \tag{5.1}$$

Here, the indices p and q sum over the set $\{0, ..., P-1\}$, representing the various energy levels. Additionally, the index σ sums over the set $\{-, +\}$, representing the spin of each fermion. The coefficient d_p represents the single particle energy corresponding to energy level p. The coefficients g^p_q are the so-called pairing strengths which represent the energy

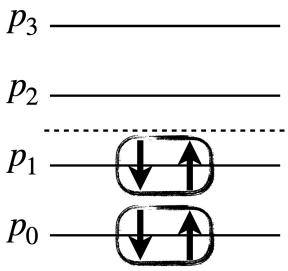


Figure 5.1: Example schematic of the paring model with $P = 4$ energy levels and $N = 2$ pairs of fermions. Shown are four energy levels with single-particle energies d_0, d_1, d_2, d_3 of which the bottom two are initially filled by pairs of fermions. The dashed line represents the Fermi level which divides the energy levels with single-particle energies d_0 and d_1 (the hole states) from those with d_2 and d_3 (the particle states).

associated with moving a pair of fermions from the q^{th} to the p^{th} energy level. We will consider various sets of pairing strengths in this section. An example of the model is represented schematically in Figure 5.1.

To simplify the pairing model, its Hamiltonian can be rewritten in terms of pairing operators as

$$H_p = \sum_{p=1}^{n} d_p N_p + \sum_{p,q=1}^{n} g_q^p A_p^\dagger A_q. \tag{5.2}$$

Here, N_p is the pair number operator which counts the number of fermions occupying the p^{th} energy level. Furthermore, $A_p^\dagger$ and A_p are the pair fermionic creation and annihilation operators, respectively, which create and annihilate pairs of fermions on the p^{th} energy level. These operators are defined in terms of fermionic creation and annihilation operators as follows

$$N_p = \sum_{\sigma} a_{p\sigma}^\dagger a_{p\sigma} \tag{5.3}$$

$$A_p^\dagger = a_{p+}^\dagger a_{p-}^\dagger \tag{5.4}$$

$$A_p = a_{p-} a_{p+}, \tag{5.5}$$

where σ sums over the set $\{+,-\}$. The purpose of this rewriting becomes clear once one notices that these operators satisfy the SU(2) algebra described by the following commutation relations

$$[A_p, A_q^\dagger] = \delta_{pq}(1 - N_p), \tag{5.6}$$

$$[N_p, A_q^\dagger] = 2\delta_{pq}A_p^\dagger, \tag{5.7}$$

$$[N_p, A_q] = -2\delta_{pq}A_p. \tag{5.8}$$

which are proven in Appendix F.

5.2 Classical Solutions

Here we solve the pairing model via various classical techniques against which we will benchmark our quantum techniques.

5.2.1 Exact Solution

If one restricts the pairing strength coefficients of the pairing model Hamiltonian (5.2) to be constant

$$H = \sum_{p=1}^{P} d_p N_p + g \sum_{p,q=1}^{P} A_p^\dagger A_q, \tag{5.9}$$

then there exists an exact solution, discovered by R.W. Richardson in 1963 [1]. The ansatz that solves the model with N pairs is given by

$$|\Psi\rangle = \prod_{\alpha=1}^{N} B_\alpha^\dagger |0\rangle, \tag{5.10}$$

where

$$B_\alpha^\dagger = \sum_{\kappa=1}^{P} \frac{1}{2d_\kappa - E_\alpha} A_\kappa^\dagger, \tag{5.11}$$

which, when plugged into the Schrodinger equation leads to a set of equations (the Richardson equations) which we shall re-derive here [46]. First, plugging the ansatz 5.10 into the Schrodinger equation and rewriting, yields

$$\begin{aligned} E\,|\Psi\rangle &= H\,|\Psi\rangle \\ &= H\prod_{\alpha=1}^{N} B_\alpha^\dagger\,|0\rangle \\ &= [H,\prod_{\alpha=1}^{N} B_\alpha^\dagger]\;|0\rangle\,, \end{aligned} \tag{5.12}$$

since $H\,|0\rangle = 0$. The commutator can be expanded as follows

$$[H,\prod_{\alpha=1}^{N} B_\alpha^\dagger] \;=\sum_{\alpha=1}^{N}\left\{\left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger\right)[H,B_\alpha^\dagger]\left(\prod_{\gamma=\alpha+1}^{N} B_\gamma^\dagger\right)\right\}, \tag{5.13}$$

the inner commutator of which is given by

$$[H,B_\alpha^\dagger] \;=[\sum_{p=1}^{P} d_p N_p + g\sum_{p,q=1}^{P} A_p^\dagger A_q,\sum_{\kappa=1}^{P}\frac{1}{2d_\kappa - E_\alpha}A_\kappa^\dagger] \tag{5.14}$$

$$=\sum_{\kappa=1}^{P}\frac{1}{2d_\kappa - E_\alpha}\left(\sum_{p=1}^{P} d_p[N_p,A_\kappa^\dagger] \;+g\sum_{p,q=1}^{P}[A_p^\dagger A_q,A_\kappa^\dagger]\right) \tag{5.15}$$

$$=\sum_{\kappa=1}^{P}\frac{1}{2d_\kappa - E_\alpha}\left(\sum_{p=1}^{P} 2\delta_{p\kappa}d_p A_p^\dagger + g\sum_{p,q=1}^{P}\delta_{q\kappa}A_p^\dagger(1-N_q)\right) \tag{5.16}$$

$$=\sum_{\kappa=1}^{P}\frac{1}{2d_\kappa - E_\alpha}\left(2d_\kappa A_\kappa^\dagger + g\sum_{p=1}^{P} A_p^\dagger(1-N_\kappa)\right) \tag{5.17}$$

$$=\sum_{\kappa=1}^{P}\left[\left(\frac{E_\alpha}{2d_\kappa - E_\alpha}+1\right)A_\kappa^\dagger + \frac{g}{2d_\kappa - E_\alpha}\sum_{p=1}^{P} A_p^\dagger(1-N_\kappa)\right] \tag{5.18}$$

$$= E_\alpha B_\alpha^\dagger + \sum_{p=1}^{P} A_p^\dagger\left(1+g\sum_{\kappa=1}^{P}\frac{1-N_\kappa}{2d_\kappa - E_\alpha}\right). \tag{5.19}$$

Plugging this back into the Schrodinger equation (5.12) and applying it to the vacuum yields

$$
\begin{aligned}
E\,|\Psi\rangle &= [H, \prod_{\alpha=1}^{N} B_\alpha^\dagger]\;|0\rangle \\
&= \sum_{\alpha=1}^{N}\left\{\left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger\right)\left[E_\alpha B_\alpha^\dagger + \sum_{p=1}^{P} A_p^\dagger\left(1 + g\sum_{\kappa=1}^{P}\frac{1-N_\kappa}{2d_\kappa - E_\alpha}\right)\right]\left(\prod_{\gamma=\alpha+1}^{N} B_\gamma^\dagger\right)\right\}|0\rangle \\
&= E\,|\Psi\rangle \\
&+ \sum_{\alpha=1}^{N}\left\{\left[\left(1 + \sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}\right)\sum_{p=1}^{P} A_p^\dagger\right]\left(\prod_{\beta=1,\beta\neq\alpha}^{N} B_\beta^\dagger\right)\right\}|0\rangle \\
&- \sum_{\alpha=1}^{N}\left\{\left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger\right)\left[\left(\sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}N_\kappa\right)\sum_{p=1}^{P} A_p^\dagger\right]\left(\prod_{\gamma=\alpha+1}^{N} B_\gamma^\dagger\right)\right\}|0\rangle\,, \qquad (5.20)
\end{aligned}
$$

where we've defined $E = \sum_{\alpha=1}^{n} E_\alpha$ and used the definition of $|\Psi\rangle$ (5.10) to simplify the first term, which implies that the Schrodinger equation is satisfied if

$$
\begin{aligned}
&\sum_{\alpha=1}^{N}\left\{\left(1 + \sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}\right)\prod_{\beta=1,\beta\neq\alpha}^{N} B_\beta^\dagger\right\}|0\rangle \\
&- \sum_{\alpha=1}^{N}\left\{\left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger\right)\left(\sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}N_\kappa\right)\left(\prod_{\gamma=\alpha+1}^{N} B_\gamma^\dagger\right)\right\}|0\rangle = 0, \qquad (5.21)
\end{aligned}
$$

where we've divided through by the constant term $\sum_p A_p^\dagger$. Note that the second term from above can be re-written as

$$
\sum_{\alpha=1}^{N}\left\{\left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger\right)\left(\sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}\right)[N_\kappa, \prod_{\gamma=\alpha+1}^{N} B_\gamma^\dagger]\right\}|0\rangle\,, \qquad (5.22)
$$

since $N_\kappa\,|0\rangle = 0$. We'll now expand the commutator from the above expression

$$
[N_\kappa, \prod_{\gamma=\alpha+1}^{N} B_\gamma^\dagger] \;=\; \sum_{\gamma=\alpha+1}^{N}\left\{\left(\prod_{\mu=\alpha+1}^{\gamma-1} B_\mu^\dagger\right)[N_\kappa, B_\gamma^\dagger]\left(\prod_{\nu=\gamma+1}^{N} B_\nu^\dagger\right)\right\}, \qquad (5.23)
$$

the inner commutator of which is

$$[N_\kappa, B_\gamma^\dagger] = \sum_{\lambda=1}^{P} \frac{1}{2d_\lambda - E_\gamma} [N_\kappa, A_\lambda^\dagger] \tag{5.24}$$

$$= \sum_{\lambda=1}^{P} \frac{2}{2d_\lambda - E_\gamma} \delta_{\kappa\lambda} A_\kappa^\dagger \tag{5.25}$$

$$= \frac{2}{2d_\kappa - E_\gamma} A_\kappa^\dagger. \tag{5.26}$$

Plugging this back into the second term (5.22) yields

$$\sum_{\alpha=1}^{N} \left\{ \left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger \right) \sum_{\gamma=\alpha+1}^{N} \left\{ \left(\prod_{\mu=\alpha+1}^{\gamma-1} B_\mu^\dagger \right) \left(\sum_{\kappa=1}^{P} \frac{2g}{(2d_\kappa - E_\alpha)(2d_\kappa - E_\gamma)} A_\kappa^\dagger \right) \left(\prod_{\nu=\gamma+1}^{N} B_\nu^\dagger \right) \right\} \right\} |0\rangle \,. \tag{5.27}$$

Applying partial fraction decomposition to the inner sum

$$\frac{1}{(2d_\kappa - E_\alpha)(2d_\kappa - E_\gamma)} = \frac{1}{E_\alpha - E_\gamma} \left(\frac{1}{(2d_\kappa - E_\alpha)} - \frac{1}{(2d_\kappa - E_\gamma)} \right), \tag{5.28}$$

turns the second term (5.27) into

$$\sum_{\alpha=1}^{N} \left\{ \left(\prod_{\beta=1}^{\alpha-1} B_\beta^\dagger \right) \sum_{\gamma=\alpha+1}^{N} \left\{ \left(\prod_{\mu=\alpha+1}^{\gamma-1} B_\mu^\dagger \right) \left[\frac{2g}{E_\alpha - E_\gamma} (B_\alpha^\dagger - B_\gamma^\dagger) \right] \left(\prod_{\nu=\gamma+1}^{N} B_\nu^\dagger \right) \right\} \right\} |0\rangle \,, \tag{5.29}$$

which can be written as

$$\sum_{\alpha=1}^{N} \sum_{\gamma=\alpha+1}^{N} \left\{ \left(\frac{2g}{E_\alpha - E_\gamma} \right) \left(\prod_{\beta=1, \beta\neq\gamma}^{N} B_\nu^\dagger \right) \right\} |0\rangle$$

$$- \sum_{\alpha=1}^{N} \sum_{\gamma=\alpha+1}^{N} \left\{ \left(\frac{2g}{E_\alpha - E_\gamma} \right) \left(\prod_{\beta=1, \beta\neq\alpha}^{N} B_\nu^\dagger \right) \right\} |0\rangle \,. \tag{5.30}$$

Switching the order of summation of the first term and swapping indices $\alpha \leftrightarrow \gamma$ followed by merging sums yields

$$- \sum_{\alpha=1}^{N} \sum_{\gamma=1}^{\alpha-1} \left\{ \left(\frac{2g}{E_\alpha - E_\gamma} \right) \left(\prod_{\beta=1, \beta\neq\alpha}^{N} B_\nu^\dagger \right) \right\} |0\rangle$$

$$-\sum_{\alpha=1}^{N}\sum_{\gamma=\alpha+1}^{N}\left\{\left(\frac{2g}{E_\alpha - E_\gamma}\right)\left(\prod_{\beta=1,\beta\neq\alpha}^{N} B_\nu^\dagger\right)\right\}|0\rangle \tag{5.31}$$

$$=-\sum_{\alpha=1}^{N}\sum_{\gamma=1,\gamma\neq\alpha}^{N}\left\{\left(\frac{2g}{E_\alpha - E_\gamma}\right)\left(\prod_{\beta=1,\beta\neq\alpha}^{N} B_\nu^\dagger\right)\right\}|0\rangle\,. \tag{5.32}$$

Finally, plugging this back into the condition that satisfies the Schrodinger equation (5.21) yields

$$\sum_{\alpha=1}^{N}\left\{\left(1+\sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}+\sum_{\gamma=1,\gamma\neq\alpha}^{N}\frac{2g}{E_\alpha - E_\gamma}\right)\left(\prod_{\beta=1,\beta\neq\alpha}^{N} B_\beta^\dagger\right)\right\}|0\rangle = 0, \tag{5.33}$$

which yields the Richardson equations

$$1+\sum_{\kappa=1}^{P}\frac{g}{2d_\kappa - E_\alpha}+\sum_{\beta=1,\beta\neq\alpha}^{N}\frac{2g}{E_\alpha - E_\beta} = 0, \tag{5.34}$$

where we've relabeled $\gamma \to \beta$. This is a set of coupled, non-linear equations from which one solves for the terms E_α and sums them to find the energy; recall

$$E = \sum_{\alpha=1}^{N} E_\alpha. \tag{5.35}$$

However, the Richardson equations are notoriously difficult to solve, due to the presence of singularities. Additionally, the pairing model can be solved exactly [60] if the single-particle energies are degenerate ($d_p = d$ for all p) and the pairing strength is separable ($g_q^p = g^p g_q$ for all p and q); that is

$$H = d\sum_{p=1}^{P} N_p + \sum_{p,q=1}^{P} g^p g_q A_p^\dagger A_q \tag{5.36}$$

However, no exact solution has been discovered for the pairing model with both arbitrary, non-degenerate single-particle orbits d_p and arbitrary pairing strengths g_q^p (5.2), save for the case ($P = 2$) with only two-energy levels [5]. This lack of an exact solution for the arbitrary case is what motivates the usage of computational techniques.

5.2.2 Full Configuration Interaction

The pairing model can be solved through exact diagonalization. In this method, the pairing Hamiltonian written in terms of pair fermionic operators (5.2) is represented as a matrix with elements

$$\begin{aligned} H_{p_i p_j, q_i q_j} &= \langle \Phi_{q_i q_j} | H | \Phi_{p_i p_j} \rangle \\ &= \delta_{p_i q_i} \delta_{p_j q_j} [2(p_i + p_j) + g^{p_i}_{p_j}] \\ &+ \delta_{p_i q_i} (1 - \delta_{p_j q_j}) g^{p_j}_{q_j} + (1 - \delta_{p_i q_i}) \delta_{p_j q_j} g^{p_i}_{q_i}, \end{aligned} \tag{5.37}$$

where $|\Phi_{p_i p_j}\rangle = a^\dagger_{p_i} a^\dagger_{p_j} |0\rangle$. Here, $i > j$ and $|0\rangle$ is the true vacuum. This Hamiltonian matrix is then diagonalized, its eigenvalues equal to the possible energies of the system.

5.2.3 Pair CCD

In this section, we apply the method of pair coupled cluster doubles theory (pCCD) to the pairing model. The pCCD equations (2.161 and 2.162) for the pairing model Hamiltonian 5.2 become

$$E = \langle \Phi_0 | \overline{H}_p | \Phi_0 \rangle , \tag{5.38}$$

$$0 = \langle \Phi^a_i | \overline{H}_p | \Phi_0 \rangle , \tag{5.39}$$

where the similarity transformed pairing Hamiltonian is

$$\overline{H}_p = e^{-T_p} H_p e^{T_p}, \tag{5.40}$$

which can be expanded via the BCH identity as

$$\overline{H}_p = H_p + \big[H_p, T_p\big] + \frac{1}{2}\big[\big[H_p, T_p\big], T_p\big] + \cdots . \tag{5.41}$$

Though this expression is infinite, it can, in this case, be truncated to just the first three terms. This is because, after the expanded form of the similarity transformed pairing Hamiltonian (5.41) is inserted into the pCCD equations (5.38 and 5.39), one can truncate the resulting expressions by noting that only certain terms in the infinite sum (those which can be fully contracted) are non-zero. The truncated pCCD equations are

$$E = \langle \Phi | \, F_p + V_p + V_p T_p \, | \Phi \rangle \,, \tag{5.42}$$

$$0 = \langle \Phi_i^a | \, V_p + F_p T_p + V_p T_p + \frac{1}{2} V_p T_p^2 - T_p V_p T_p \, | \Phi \rangle \,, \tag{5.43}$$

where

$$F_p = \sum_p d_p N_p, \tag{5.44}$$

$$V_p = \sum_{pq} g_q^p A_p^\dagger A_q, \tag{5.45}$$

with the partitioning of the pairing model Hamiltonian a single-body and two-body term: $H_p = F_p + V_p$. We can save ourselves some work by recognizing that the first two terms of the truncated pCCD energy equation (5.42) equal the reference energy, and hence the pCCD correlation energy equation is

$$\Delta E = \langle \Phi | V_p T_p | \Phi \rangle \,. \tag{5.46}$$

Using Wick's theorem, we calculate the truncated pCCD energy equation (5.42)

$$\begin{aligned} \langle \Phi | V_p T_p | \Phi \rangle &= \sum_{pqia} g_q^p t_i^a \, \langle \Phi | A_p^\dagger A_q A_a^\dagger A_i | \Phi \rangle \\ &= \sum_{pqia} g_q^p t_i^a \, \langle \Phi | \overbrace{a_{p+}^\dagger a_{p-}^\dagger a_{q-} a_{q+} a_{a+}^\dagger a_{a-}^\dagger a_{i-} a_{i+}}^{\text{}} | \Phi \rangle \end{aligned}$$

$$= \sum_{pqia} g_q^p t_i^a \delta_{pi} \delta_{qa}$$

$$= \sum_{ia} g_a^i t_i^a . \tag{5.47}$$

Turning our attention now to the truncated pCCD amplitude equation (5.42), we start with

$$\langle \Phi_i^a | V_p | \Phi \rangle = \sum_{pq} g_q^p \langle \Phi | A_i^\dagger A_a A_p^\dagger A_q | \Phi \rangle$$

$$= \sum_{pq} g_q^p \langle \Phi | \overbrace{a_{i+}^\dagger \overbrace{a_{i-}^\dagger \overbrace{a_{a-} \overbrace{a_{a+} a_{p+}^\dagger}^{} a_{p-}^\dagger}^{} a_{q-}}^{} a_{q+}}^{} | \Phi \rangle$$

$$= \sum_{pq} g_q^p \delta_{pa} \delta_{qi}$$

$$= g_i^a . \tag{5.48}$$

To compute the next term $\langle \Phi_i^a | F_p T_p | \Phi \rangle$, instead of writing out all possible contractions, it will be easier to compute the un-truncated term from which the this term comes, namely $\langle \Phi_i^a | [F_p, T_p] | \Phi \rangle$. We do so by using the commutation relations (5.7 and 5.8) as follows

$$\langle \Phi_i^a | [F_p, T_p] | \Phi \rangle = \sum_{pjb} d_p t_j^b \langle \Phi_i^a | \, [N_p, A_b^\dagger A_j] \, | \Phi \rangle$$

$$= \sum_{pjb} d_p t_j^b \langle \Phi_i^a | \left([N_p, A_b^\dagger] \, A_j - A_b^\dagger [N_p, A_j] \right) | \Phi \rangle$$

$$= 2 \sum_{pjb} d_p t_j^b (\delta_{pb} - \delta_{pj}) \langle \Phi_i^a | \, A_b^\dagger A_j \, | \Phi \rangle$$

$$= 2 \sum_{jb} (d_b - d_j) t_j^b \langle \Phi | \, A_i^\dagger A_a A_b^\dagger A_j \, | \Phi \rangle$$

$$= 2 \sum_{jb} (d_b - d_j) t_j^b \langle \Phi | \overbrace{a_{i+}^\dagger \overbrace{a_{i-}^\dagger \overbrace{a_{a-} \overbrace{a_{a+} a_{b+}^\dagger}^{} a_{b-}^\dagger}^{} a_{j-}}^{} a_{j+}}^{} | \Phi \rangle$$

$$= 2 \sum_{jb} (d_b - d_j) t_j^b \delta_{ij} \delta_{ab}$$

$$= 2(d_a - d_i)t_i^a. \tag{5.49}$$

The same is true for the next term $\langle\Phi_i^a|V_pT_p|\Phi\rangle$ and so it will be calculated analogously

$$\begin{aligned}
\langle\Phi_i^a|[V_p, T_p]\,|\Phi\rangle &= \sum_{pqjb} g_q^p t_j^b\, \langle\Phi_i^a|[A_p^\dagger A_q, A_b^\dagger A_j]\,|\Phi\rangle \\
&= \sum_{pqjb} g_q^p t_j^b\, \langle\Phi_i^a|\,(A_p^\dagger \cancel{A_b^\dagger[A_q, A_j]} + A_p^\dagger[A_q, A_b^\dagger]\,A_j \\
&\qquad\qquad + A_b^\dagger[A_p^\dagger, A_j]\,A_q + \cancel{[A_p^\dagger, A_b^\dagger]}\,A_jA_q)\,|\Phi\rangle \\
&= \sum_{pqjb} g_q^p t_j^b\, \langle\Phi_i^a|\,(\delta_{qb}A_p^\dagger(1-N_b)A_j + \delta_{pj}A_b^\dagger(N_j-1)A_q)\,|\Phi\rangle \\
&= \sum_{pqjb} g_q^p t_j^b\, \langle\Phi_i^a|\,(\delta_{qb}A_p^\dagger(1-N_b)A_j + \delta_{pj}A_b^\dagger A_q(N_j-1) - 2\delta_{pj}\delta_{qj}A_b^\dagger A_j)\,|\Phi\rangle \\
&= \sum_{jb} t_j^b\, \langle\Phi|\,A_i^\dagger A_a(g_b^p A_p^\dagger A_j + g_q^j A_b^\dagger A_q - 2g_j^j A_b^\dagger A_j)\,|\Phi\rangle \\
&= \sum_{jb} t_j^b(g_b^p\, \langle\Phi|\overbrace{a_{i+}^\dagger a_{i-}^\dagger a_{a-}a_{a+}a_{p+}^\dagger a_{p-}^\dagger a_{j-}a_{j+}}|\Phi\rangle \\
&\qquad + g_q^j\, \langle\Phi|\overbrace{a_{i+}^\dagger a_{i-}^\dagger a_{a-}a_{a+}a_{b+}^\dagger a_{b-}^\dagger a_{q-}a_{q+}}|\Phi\rangle \\
&\qquad - 2g_j^j\, \langle\Phi|\overbrace{a_{i+}^\dagger a_{i-}^\dagger a_{a-}a_{a+}a_{b+}^\dagger a_{b-}^\dagger a_{j-}a_{j+}}|\Phi\rangle \\
&= \sum_{jb} t_j^b(g_b^p\delta_{ij}\delta_{ap} + g_q^j\delta_{iq}\delta_{ab} - 2g_j^j\delta_{ij}\delta_{ab}) \\
&= \sum_b g_b^a t_i^b + \sum_j g_i^j t_j^a - 2g_i^i t_i^a,
\end{aligned} \tag{5.50}$$

where we've used the facts that $[N_b, A_j] = 0$, $[N_j, A_q] = -2\delta_{qj}A_j$, and $(1-N_b)\,|\Phi\rangle = (N_j-1)\,|\Phi\rangle = 1$. We will skip the penultimate term for now, the reasons for which will

become clear later, and compute the final term first

$$
\begin{aligned}
&- \left\langle \Phi_i^a \middle| T_p V_p T_p \middle| \Phi \right\rangle \qquad (5.51)\\
&= - \sum_{pqjkbc} g_q^p t_j^b t_k^c \, \langle \Phi | A_i^\dagger A_a A_b^\dagger A_j A_p^\dagger A_q A_c^\dagger A_k | \Phi \rangle \\
&= - \sum_{pqjkbc} g_q^p t_j^b t_k^c \, \langle \Phi | \overbrace{a_{i+}^\dagger \overbrace{a_{i-}^\dagger \overbrace{a_{a-} \overbrace{a_{a+}^\dagger a_{b+}^\dagger}^{} a_{b-}^\dagger}^{} a_{j-}}^{} a_{j+}}^{} \overbrace{a_{p+}^\dagger \overbrace{a_{p-}^\dagger \overbrace{a_{q-} \overbrace{a_{q+}^\dagger a_{c+}^\dagger}^{} a_{c-}^\dagger}^{} a_{k-}}^{} a_{k+}}^{} | \Phi \rangle \\
&= - \sum_{pqjkbc} g_q^p t_j^b t_k^c \delta_{ij} \delta_{ab} \delta_{pk} \delta_{qc} \\
&= - \sum_{kc} g_c^k t_i^a t_k^c \\
&= - \sum_{jb} g_b^j t_i^a t_j^b, \qquad (5.52)
\end{aligned}
$$

where we've relabeled $k \to j$ and $c \to b$. We now come to the final term $\langle \Phi_i^a | V_p T_p^2 | \Phi \rangle$, which we left for last as it is most easily solved through diagrammatic methods due to the large number of possible contractions. The term is calculated up to the point of contractions as

$$
\begin{aligned}
&\frac{1}{2} \langle \Phi_i^a | \, V_p T_p^2 \, | \Phi \rangle \\
&= \frac{1}{2} \sum_{pqjkbc} g_q^p t_j^b t_k^c \, \langle \Phi | \, A_i^\dagger A_a A_p^\dagger A_q A_b^\dagger A_j A_c^\dagger A_k \, | \Phi \rangle \\
&= \frac{1}{2} \sum_{pqjkbc} g_q^p t_j^b t_k^c \, \langle \Phi | \, a_{i+}^\dagger a_{i-}^\dagger a_{a-} a_{a+} a_{p+}^\dagger a_{p-}^\dagger a_{q-} a_{q+} a_{b+}^\dagger a_{b-}^\dagger a_{j-} a_{j+} a_{c+}^\dagger a_{c-}^\dagger a_{k-} a_{k+} \, | \Phi \rangle \, ,
\end{aligned} \qquad (5.53)
$$

at which point we use a novel extension of Goldstone diagrams (which I've named pair-Goldstone diagrams) to continue.

In standard Goldstone diagrams, vertices represent macro-operators (like F_p, V_p, and T as defined above) while lines between vertices represent contractions between fermionic

operators contained within the macro-operators represented by said vertices. Lines directed to the right or left represent contractions between hole and particle operators, respectively. The resulting expressions can be read directly from the diagrams by identifying lines entering or leaving a vertex as the lower or upper index of the prefactor represented by said vertex, respectively.

However, standard Goldstone diagrams do not visually capture additional Kronecker delta's that can be created when dealing with pairing Hamiltonians. To demonstrate, consider Figure 5.2a which is a traditional Goldstone diagram that represents one possible set of contractions that result from (5.53) which comes from the term $\langle \Phi_i^a | V_p T_p^2 | \Phi \rangle$. In this diagram, the top left vertex represents the excitation operator $E_a^i = A_i^\dagger A_a$ (which creates the excited state $\langle \Phi | E_a^i = \langle \Phi_i^a |$), the bottom left vertex represents V_p, and the top and bottom right diagrams represent the two T_p diagrams. This particular diagram represents the contractions that result in the product

$$\delta_{i+j+}\delta_{a+b+}\delta_{a-b-}\delta_{i-k-}\delta_{p-j-}\delta_{q-c-}\delta_{q+c+}\delta_{p+k+}, \tag{5.54}$$

which results in the following transformation

$$\sum_{pqjkbc} g_q^p t_j^b t_k^c \rightarrow \sum_{jc} g_c^i t_i^a t_i^c, \tag{5.55}$$

which is equivalent, upon relabeling $c \rightarrow b$, to

$$\sum_{jb} g_b^i t_i^a t_i^b, \tag{5.56}$$

which is immediately read from the diagram when labeling the rightward facing arrows (from top to bottom) $i+$, $i-$, $j-$, and $j+$ and the leftward facing arrows (from top to bottom) $a+$, $a-$, $b-$, and $b+$.

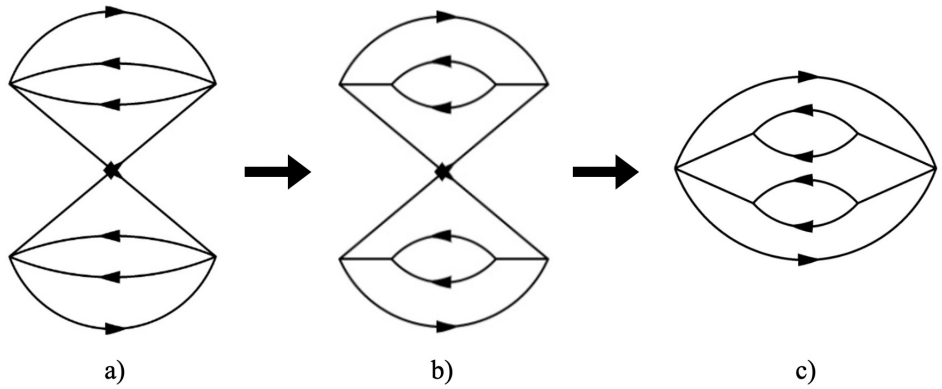


Figure 5.2: Pair-Goldstone contraction schematic.

Note that we actually reduced the number of summed indices by three despite only having eight contractions (which would normally result in a reduction of 2 indices). This occurs because we are working with a pair operators. To see this, note that the products $\delta_{i+j+}\delta_{i-k-}$ and $\delta_{p-j-}\delta_{p+k+}$ imply $\delta_{pi}\delta_{ji}\delta_{ki}$. That is, the indices p, j, and k, all go to i. This "additional" contraction is not immediately obvious from the traditional Goldstone diagram (Figure 5.2a). However, it is in the pair-Goldstone diagram representation (Figures 5.2b and 5.2c). Here, each vertex is replaced with a solid line, one end of which is exclusively for particle lines while the other is exclusively for hole lines. There will always be two lines attached to each end; one for the + state of the pair and the other for the − state. This allows one to visually capture the phenomenon of pairs of fermionic operators contracting to operators with differing indices. Notice that in Figure 5.2b, the hole lines of the top T_p operator are $i+$ and $j-$, while the hole lines of the bottom T_p operator are $i-$ and $j+$. This clearly implies that the bottom indices of both t terms should be i. This is visually shown in the "contracted diagram" Figure 5.2c, in which the hole ends of the T_p lines have been brought or "contracted" together. This additionally results in the merging of the $i+$ and $j-$ lines, as well as the $i-$ and $j+$ lines. The lines are now labeled (from top to bottom)

$i+$, $a+$, $a-$, $b-$, $b+$, $i-$. Note that because of all this, the final expression (5.56) can be immediately read from the pair-Goldstone representation (Figure 5.2c).

We now use pair-Goldstone diagrams to calculate the term $\langle \Phi_i^a | \, V_p T_p^2 \, | \Phi \rangle$ (5.53). That is, the pair-Goldstone diagrams from Figure 5.3 result in the following expressions

$$a) \rightarrow 2 \sum_{bj} g_b^j t_j^b t_i^a \tag{5.57}$$

$$b) \rightarrow 2 \sum_{bj} g_b^j t_j^a t_i^b \tag{5.58}$$

$$c) \rightarrow -4 \sum_{b} g_b^i t_i^b t_i^a \tag{5.59}$$

$$d) \rightarrow -4 \sum_{j} g_a^j t_j^a t_i^a \tag{5.60}$$

$$e) \rightarrow 4 g_a^i t_i^a t_i^a . \tag{5.61}$$

The minus signs in front of the expressions resulting from diagrams c) and d) come from the fact that their pre-contracted diagrams (see for example Figure 5.2b) contain an odd number of crossings. This can be immediately obtained from the contracted diagrams by counting the number of pairs of ends of lines that have been merged; in diagrams c) and d), one pair of ends has been merged (either the hole ends or the particle ends) resulting in a prefactor sign of $(-1)^1$, while in diagram e), two pairs of ends have been merged (both the hole ends and the particle ends) resulting in a prefactor sign of $(-1)^2$. The pre-factors of the expressions resulting from diagrams a) and b) can be determined by counting the number of symmetries in each diagram (the same as in standard Goldstone diagrams) for which each is 2. The pre-factors of the contracted diagrams c), d), and e) can be determined by counting the number of unique ways that each of these diagrams can be created by contracting the un-contracted diagrams, a) and b), for which each is four (two from each

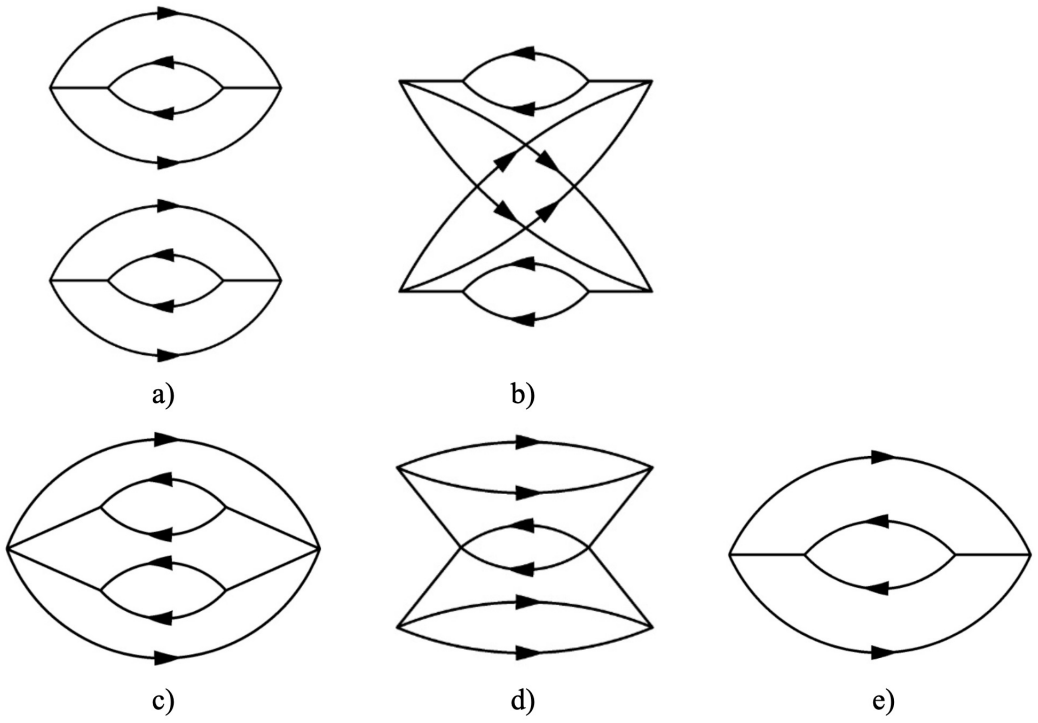


Figure 5.3: Pair-Goldstone diagrams for the pairing model.

un-contracted diagram).

Plugging these resulting expression (5.57)-(5.61) into equation (5.53) gives us

$$\frac{1}{2}\langle\Phi_i^a|\,V_pT_p^2\,|\Phi\rangle = \sum_{bj} g_b^j t_j^b t_i^a + \sum_{bj} g_b^j t_j^a t_i^b - 2\sum_{b} g_b^i t_i^b t_i^a - 2\sum_{j} g_a^j t_j^a t_i^a + 2g_a^i t_i^a t_i^a. \tag{5.62}$$

All together, the pCCD correlation energy equation (5.46) becomes

$$\Delta E = \sum_{ia} g_a^i t_i^a, \tag{5.63}$$

while the pCCD amplitude equations (5.43) become

$$0 = g_i^a + 2\Big(d_a - d_i - g_i^i + g_a^i t_i^a - \sum_b g_b^i t_i^b - \sum_j g_a^j t_j^a\Big)t_i^a + \sum_b g_b^a t_i^b + \sum_j g_i^j t_j^a + \sum_{bj} g_b^j t_j^a t_i^b. \tag{5.64}$$

In practice, the amplitudes t are solved for by solving the amplitude equation (5.64), which is non-linear and coupled, via an iterative root finding algorithm such as Newton's method and plugging them into the energy equation (5.63).

Although we were able to skip over deriving the reference energy of the pairing model in the previous section because we only cared about the correlation energy ΔE, we will need an expression for the reference energy when we turn to implementing unitary pair coupled cluster (UpCC) theory as the ansatz for the variational quantum eigensolver (VQE). Thus, we derive said expression here using Wick's theorem as here

$$\begin{aligned}
E_{\text{ref}} &= \langle\Phi|H_p|\Phi\rangle \\
&= \langle\Phi|F_p|\Phi\rangle + \langle\Phi|V_p|\Phi\rangle \\
&= \sum_{p\sigma} d_p \langle\Phi|\overbrace{a^\dagger_{p\sigma} a_{p\sigma}}|\Phi\rangle + \sum_{pq} g^p_q \langle\Phi|\overbrace{a^\dagger_{p+} \overbrace{a^\dagger_{p-} a_{q-}} a_{q+}}|\Phi\rangle \\
&= 2\sum_p d_p h(p) + \sum_{pq} g^p_q h(p)\delta_{pq} \\
&= \sum_i (2d_i + g^i_i). \qquad (5.65)
\end{aligned}$$

In order to find the amplitudes t, Newton's method must converge, which requires a good initial guess for each t. To find one, we turn to many-body perturbation theory.

5.2.4 Many-body Perturbation Theory

The single particle energies ϵ_p (2.107) for the pairing Hamiltonian are given by

$$\begin{aligned}
\epsilon_p &= t^p_p + \sum_i v^{pi}_{pi} \qquad (5.66) \\
&= d_p + \sum_i \delta_{pi} g^p_p \qquad (5.67) \\
&= d_p + h(p) g^p_p, \qquad (5.68)
\end{aligned}$$

and the interacting Hamiltonian H_I is given by

$$H_I = \sum_{p=1}^{P}(d_p - \epsilon_p)N_p + \sum_{p,q=1}^{P} g_q^p A_p^\dagger A_q. \tag{5.69}$$

For the pairing model, the second and third order energy correlation contributions (2.155 - 2.156) are

$$\begin{aligned}\Delta E^{(2)} &= \sum_{ai} \frac{g_i^a g_a^i}{\epsilon_i^a} \\ &= \frac{1}{2}\sum_{ai} \frac{g_i^a g_a^i}{d_i - d_a + g_i^i},\end{aligned} \tag{5.70}$$

and

$$\begin{aligned}\Delta E^{(3)} &= \sum_{iab} \frac{g_b^i g_a^b g_i^a}{\epsilon_i^a \epsilon_i^b} + \sum_{ija} \frac{g_j^a g_i^j g_a^i}{\epsilon_i^a \epsilon_j^a} \\ &= \frac{1}{4}\sum_{iab} \frac{g_b^i g_a^b g_i^a}{\left(d_i - d_a + g_i^i\right)\left(d_i - d_b + g_i^i\right)} + \sum_{ija} \frac{g_j^a g_i^j g_a^i}{\left(d_i - d_a + g_i^i\right)\left(d_j - d_a + g_j^j\right)}.\end{aligned} \tag{5.71}$$

One good initial guess for t can be found by mathcing the pCCD correlation energy (5.63) with the second order contribution to the correlation energy from MBPT (5.70)

$$\sum_{ia} g_a^i t_i^{a(0)} = \sum_{ia} g_a^i \frac{g_i^a}{\epsilon_i^a}, \tag{5.72}$$

which allows us to identify

$$\begin{aligned}t_i^{a(0)} &= \frac{g_i^a}{\epsilon_i^a} \\ &= \frac{1}{2}\frac{g_i^a}{d_i - d_a + g_i^i},\end{aligned} \tag{5.73}$$

for the pairing model. Many-body perturbation theory provides a more tractable, yet still approximate, solution.

5.3 Quantum Solutions

5.3.1 Mapping the Hamiltonian

To solve the pairing model on a quantum computer, we must first determine a mapping of our problem to qubits. To do so, we simply let qubit p represent energy level p. In our pairing scheme, the qubit being in state $|0\rangle$ or $|1\rangle$ implies that its corresponding energy level is completely unoccupied or occupied by a pair of fermions, respectively. Note that there is no possibility for broken pairs in this mapping. Now that we've chosen a qubit mapping, we must transform the Hamiltonian accordingly, from fermionic operators to spin operators. To do so, we'll first separate the pairing Hamiltonian (5.2) into strict one-body and two-body terms, as follows

$$H = \frac{1}{2}\sum_{p=1}^{P}(2d_p + g_p^p)N_p + \sum_{\substack{p,q=1 \\ p\neq q}}^{P} g_q^p A_p^\dagger A_q, \tag{5.74}$$

which follows from the fact that, when $p = q$ in the second sum of (5.2), it results in the operator

$$A_p^\dagger A_p = \left(\frac{X_p - iY_p}{2}\right)\left(\frac{X_p + iY_p}{2}\right) \tag{5.75}$$

$$= \frac{I_p - Z_p}{2} \tag{5.76}$$

$$= \frac{N_p}{2}. \tag{5.77}$$

Using the pair-fermionic anti-commutation relation (5.6) and applying the Jordan-Wigner transformation to (5.77) yields

$$\frac{N_p}{2} = \frac{1}{2}\left(I_p - \delta_{pq}[A_p, A_q^\dagger]\right) \tag{5.78}$$

$$= \frac{1}{2}\left(I_p - \frac{1}{4}\delta_{pq}[X_p + iY_p, X_q - iY_q]\right) \tag{5.79}$$

$$= \frac{I_p - Z_p}{2}. \tag{5.80}$$

To deal with the second term of (5.74), we note that the sum can be broken up as

$$\sum_{\substack{p,q=1 \\ p\neq q}}^{P} g_q^p A_p^\dagger A_q = \sum_{\substack{p,q=1 \\ p<q}}^{P} g_q^p A_p^\dagger A_q + \sum_{\substack{p,q=1 \\ q<p}}^{P} g_q^p A_p^\dagger A_q \tag{5.81}$$

$$= \sum_{\substack{p,q=1 \\ p<q}}^{P} g_q^p A_p^\dagger A_q + \sum_{\substack{p,q=1 \\ p<q}}^{P} g_q^p A_q^\dagger A_p \tag{5.82}$$

$$= \sum_{\substack{p,q=1 \\ p<q}}^{P} g_q^p (A_p^\dagger A_q + A_p A_q^\dagger), \tag{5.83}$$

where we've swapped the indices $p \leftrightarrow q$ to obtain (5.82) and used the fact that $[A_p, A_q^\dagger] = 0$ for $p \neq q$ (which $p < q$ implies) to obtain (5.83). Applying the Jordan-Wigner transformation to (5.83) yields

$$A_p^\dagger A_q + A_p A_q^\dagger = \left(\frac{X_p - iY_p}{2}\right)\left(\frac{X_q + iY_q}{2}\right) + \left(\frac{X_p + iY_p}{2}\right)\left(\frac{X_q - iY_q}{2}\right) \tag{5.84}$$

$$= \frac{X_p X_q + Y_p Y_q}{2}. \tag{5.85}$$

All together, the pairing Hamiltonian (5.74), after Jordan-Wigner transformation, becomes

$$H = \frac{1}{2}\sum_{p=1}^{P}(2d_p + g_p^p)(I_p - Z_p) + \frac{1}{2}\sum_{\substack{p,q=1 \\ p<q}} g_q^p (X_p X_q + Y_p Y_q). \tag{5.86}$$

5.3.2 Mapping the Ansatz

As an ansatz for the pairing model, we choose the unitary pair coupled cluster doubles (UpCCD) ansatz, which is the unitary version of the pair coupled cluster doubles (pCCD) ansatz. It is chosen because any quantum ansatz must be unitary (as all quantum gates

must be unitary), it is a pairing ansatz, and includes only up to doubles which allows for the use of, at most, efficiently decomposable two-qubit gates. It is given by

$$|\Psi\rangle = e^{T_p - T_p^\dagger} |\Phi\rangle \,. \tag{5.87}$$

Recalling that T_p, as defined in (2.168), is given by

$$T_p = \sum_{ia} t_i^a A_a^\dagger A_i, \tag{5.88}$$

turns the ansatz into the variational form

$$|\Psi(t)\rangle = \exp\left\{ \sum_{ia} t_i^a \left(A_a^\dagger A_i - A_a A_i^\dagger \right) \right\} |\Phi_0\rangle \,, \tag{5.89}$$

where we've used the fact that $[A_a, A_i^\dagger] = 0$ for since, by definition $a \neq i$. Applying the Jordan Wigner transformation to the operators yields

$$A_a^\dagger A_i - A_a A_i^\dagger = \left(\frac{X_a - iY_a}{2} \right) \left(\frac{X_i + iY_i}{2} \right) - \left(\frac{X_a + iY_a}{2} \right) \left(\frac{X_i - iY_i}{2} \right) \tag{5.90}$$

$$= \frac{i}{2} (X_a Y_i - Y_a X_i). \tag{5.91}$$

Plugging this back into the ansatz (5.89) yields

$$|\Psi(t)\rangle = \exp\left\{ \frac{i}{2} \sum_{ia} t_i^a (X_a Y_i - Y_a X_i) \right\} |\Phi_0\rangle \,. \tag{5.92}$$

As it is not known how to efficiently decompose n-qubit operators where n is greater than two, we must write our ansatz as a sum of two-qubit operators. To achieve this, we employ a simple Suzuki-Trotter approximation [82], which we truncate to a single first-order step in order to minimize the depth of the quantum circuit that will implement it. Doing so yields

$$|\Psi(t)\rangle = \prod_{ia} A_i^a |\Phi_0\rangle \,, \tag{5.93}$$

where we've defined the two-qubit operator

$$A_i^a = \exp\left\{\frac{i}{2} t_i^a (X_a Y_i - Y_a X_i)\right\}, \tag{5.94}$$

which is efficiently decomposable. One could consider higher order Trotterizations or multiple Trotter steps; however, employing these would substantially increase the depth of the circuit. Additionally, it will be seen later that our simple Trotter approximation provides a sufficient ansatz for VQE to be able to minimize the energy. The operator A_i^a takes the following matrix form

$$A_i^a = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos t_a^i & \sin t_a^i & 0 \\ 0 & -\sin t_a^i & \cos t_a^i & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{5.95}$$

From this matrix form, one can see that A_i^a acts non-trivially only in the two-qubit subspace $\{|01\rangle, |10\rangle\}$ and is closed in said subspace. This means that as long as we initialize our quantum circuit to a bit-string state with Hamming weight equal to the number of particles, we're guaranteed (save for noise) to only search the relevant subspace of the Hamiltonian.

5.3.3 Implementing the Ansatz

Now that we've mapped the ansatz to spin operators, we must now determine how to efficiently decompose it into a quantum circuit given the limitations of qubit connectivity and circuit depth allowed on NISQ era devices. Here, we give two such methods; the first is for a quantum computer with linear connectivity while the second is for one with circular connectivity.

To define these two connectivity terms, consider a quantum computer with N qubits. Label and order the qubits as $0, 2, ..., N-1$. A quantum computer has linear connectivity

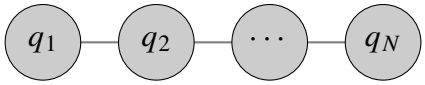


Figure 5.4: Linear qubit-connectivity schematic.

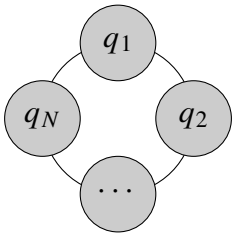
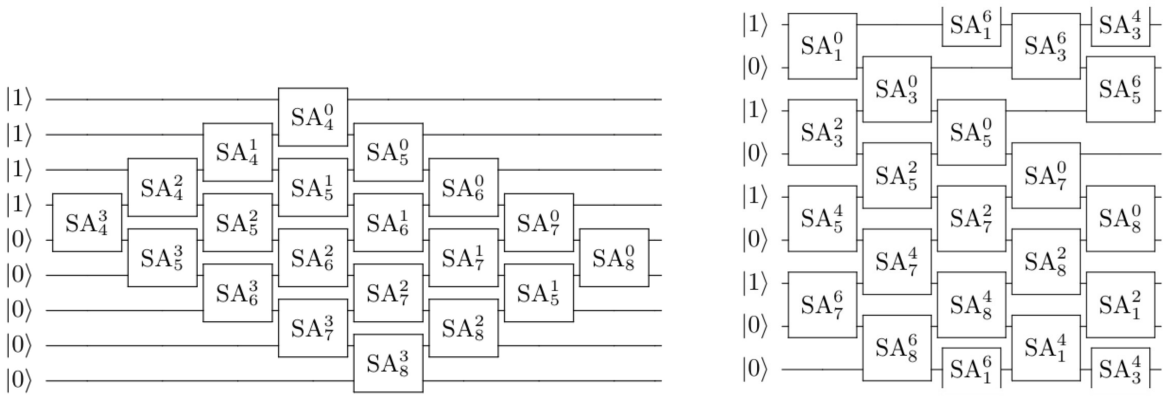


Figure 5.5: Circular qubit-connectivity schematic.

when each of its qubits, except for the first and last, is connected to its left and right neighboring qubits. The first and last qubits are connected to their only neighboring qubit. This is visualized in Figure 5.4 where the qubits are connected if the circles representing them are connected in the graph. A quantum computer has circular connectivity when each of its qubits (including the first and last) is connected to its left and right neighboring qubits. This is visualized in Figure 5.5. The ansatz is implemented using what we'll call the particle-hole swap network (phsn) technique which has both a linear connectivity [30] and a novel circular connectivity version. An illustration of the circuits for the four-particle, five-hole system for both connectivities are given in Figure 5.6.

In said circuits, S is the SWAP gate 3.37 and A_a^i is the two qubit operator defined in (5.94). Because $SA_a^i \in \mathrm{SO}(4)$, it can be efficiently mapped to a depth five circuit consisting of two CNOTs and twelve single-qubit gates [83]. With P being the number of energy levels and N being the number of energy levels that are initially filled, the number of two-qubit gates SA_a^i is equal to $G = P(P - N)$. The depth of the linear particle-hole swap network is $D(\mathrm{lphsn}) = P - 1$ where P is the number of qubits (energy-levels in the pairing model). The depth of the circular particle-hole swap network, however, is



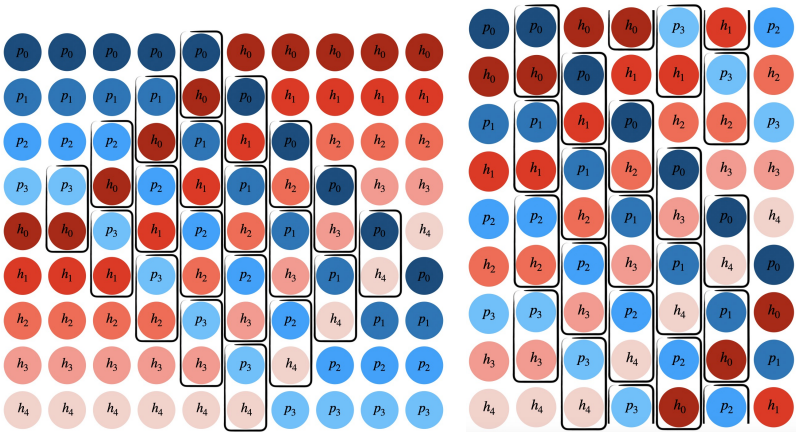
(a) (b)

Figure 5.6: a) Linear particle-hole swap network (lphsn) for a four-particle, five-hole system. b) Circular particle-hole swap network for a four-particle, five-hole system (cphsn). See Figure 5.7 for schematic representation.

$D(\text{cphsn}) = \max(N, P - N)$ where N is the number of particles (Hamming weight of the initial state). Because N is bounded as $1 \leq N \leq P$, we have that the depth of the cphsn is bounded as $\lceil P/2 \rceil \leq D(\text{cphsn}) \leq P - 1$. The depth decreases as N approached $P/2$ from either direction, achieving a minimum of $\lceil P/2 \rceil$. Thus, we have that $D(\text{cphsn}) \leq D(\text{lphsn})$. That is, the circular connectivity enables a circuit with depth less than or equal to that of a circuit constrained by linear connectivity, up to reduction by a factor of 2. Shortening the depth of the circuit is vital for decreasing noise on NISQ era devices. The algorithm to implement the linear particle-hole swap network is given below.

5.3.4 Initialization

As VQE involves a minimization algorithm, it is important that we choose a good initial guess for the variational parameters t_i^a so as to give the classical minimization algorithm the best chance of finding the ground state energy quickly and with high precision. The



(a) (b)

Figure 5.7: a) Schematic representation of the linear particle-hole swap network (lphsn) for a four-particle, five-hole system. b) Schematic representation of the circular particle-hole swap network (cphsn) for a four-particle, five-hole system. Each circle represents a qubit and a slot (particle/hole). The particles are labeled $p_0, ..., p_3$ and are colored blue while the holes are labeled $h_0, ..., h_4$ and are colored red. The first and last columns of circles are the initial and final positions of the qubits/slots. A rectangle around a pair of circles (p_i, h_j) denotes that the gate $SA^{p_i}_{h_j}$ has been applied between the corresponding qubits. See Figure 5.6 for circuit representation.

initial guess we will be using is

$$t_i^{a\,(0)} = \frac{1}{2}\frac{g_i^a}{d_i - d_a + g_i^i}, \tag{5.96}$$

from (5.73) which was computed by comparing the pCCD correlation energy with the second order contribution to the correlation energy from MBPT in subsection 5.2.4. This provides a good initial guess because we know that it results in an energy that is close to the second order contribution to the correlation energy from MBPT, which in turn is a good approximation to the true ground state. This means that our minimizer will start close to the correct solution in the energy landscape through which it must transverse. It also implies that, barring degeneracy, an ansatz initialized to $t_i^{a\,(0)}$ should have a significant overlap with the true ground state wavefunction.

Algorithm 5.1 Linear Particle-Hole Swap Network

Input: Number of energy levels p, number of initially filled energy levels n, gates SA_a^i, and qubit order list $o = [1, 2, ..., n]$.
Output: Quantum circuit implementation of lphsn.

$t = n - 1$ ▷ index of highest qubit acted upon
$b = n$ ▷ index of lowest qubit acted upon
$h = 1$ ▷ height (# of gates per column)
$d_t = -1$ ▷ top direction
$d_b = 1$ ▷ bottom direction
for $0 \leq i \leq n - 1$ **do**
    apply gate X to qubit i
end for
while $h \neq 0$ **do**
    for $0 \leq i \leq h - 1$ **do**
        $x = t + 2i$
        apply $\mathrm{SA}_{o[x+1]}^{o[x]}$ to qubits x and $x + 1$
        $o[x], o[x + 1] = o[x + 1], o[x]$
    end for
    if $t = 1$ **then**
        $d_t = 1$
    end if
    if $b = p$ **then**
        $d_b = -1$
    end if
    $t \mathrel{+}= d_t$, $b \mathrel{+}= b_d$, $h = (b - t + 1)/2$
end while

Figure 5.8 compares initial correlation energies $\langle\psi(\theta)|H|\psi(\theta)\rangle$ between the case where the initial parameters t_i^a are chosen randomly and where they are informed by many-body perturbation theory, $t_i^a = t_i^{a\,(0)}$ (5.96); It uses the pairing model with $P = 4$ energy levels and $N = 2$ pairs of particles. It can be seen from said figure that the initial correlation energy resulting from $t_i^a = t_i^{a\,(0)}$ (E_calc_ia) is much closer to the correct ground state correlation energy (E_true) than the initial correlation energy resulting from random initial

Algorithm 5.2 Circular Particle-Hole Swap Network

Input: Number of energy levels p, number of initially filled energy levels n, gates SA_a^i, and qubit order list $o = [1, 2, ..., n]$.
Output: Quantum circuit implementation of cphsn.

$d = \max(n, p - n)$ ▷ depth of circuit
$h = \min(n, p - n)$ ▷ height (# of gates per column)
for $0 \leq i \leq n - 1$ **do**
    apply gate X to qubit $2i$
end for
if $n \geq \lceil p/2 \rceil$ **then**
    for $0 \leq i \leq p - 1$ **do**
        apply gate X to qubit i
    end for
end if
for $0 \leq i \leq d - 1$ **do**
    for $0 \leq i \leq h - 1$ **do**
        $x = (i + 2j) \mod p$
        apply $\mathrm{SA}_{o[x+1]}^{o[x]}$ to qubits x and $x + 1$
        $o[x], o[x + 1] = o[x + 1], o[x]$
    end for
end for

parameters (E_calc_rand). In fact, the randomness of the initial parameters case seems to carry over to the corresponding initial correlation energy as it produces a noisy line that is way above the correct correlation energies. Here, we've defined the initial correlation energies as

$$E_{\text{calc\_ia}} = \left\langle \Phi\left(t_i^{a(0)}\right) \middle| H \middle| \Phi\left(t_i^{a(0)}\right) \right\rangle, \tag{5.97}$$

$$E_{\text{calc\_rand}} = \langle \Phi(t_{\text{rand}}) | H | \Phi(t_{\text{rand}}) \rangle, \tag{5.98}$$

$$E_{\text{true}} = \langle \Phi(t_{\text{min}}) | H | \Phi(t_{\text{min}}) \rangle, \tag{5.99}$$

where t_{rand} is a set of random parameters while t_{min} is the set of parameters that minimize the energy.

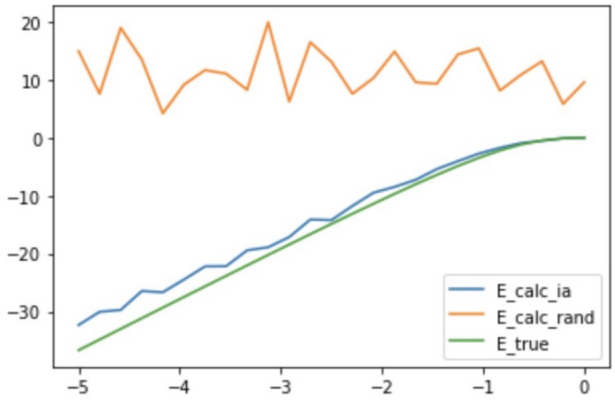


Figure 5.8: Initial correlation energies for the pairing model with $P = 4$ energy levels and $N = 2$ pairs of fermions are compared to E_true, the true ground state correlation energy. E_calc_ia uses the initial parameters informed by MBPT (5.96) and E_calc_rand uses random initial parameters.

Because (E_calc_ia) is so much closer to (E_true) than (E_calc_rand), the optimization portion of the algorithm should not have to go through as many iterations when using $t_i^a = t_i^{a\,(0)}$ as opposed to $t_i^a = t_{\text{rand}}$ which is confirmed by Figure 5.9. This illustrates the importance of initializing one's variational parameters in an informative way than just doing so randomly. This requires knowledge about the specific system that one is trying to solve. In this case, we were able to use information from many-body perturbation theory and coupled cluster theory to inform ourselves of a good initial guess.

Finally, Figure 5.10 compares the minimized correlation energy between the two initialization cases, showing that the MBPT informed (5.96) initial parameterization leads to more accurate predictions of the ground state correlation energy.

5.3.5 Results

To start, we compare the performances of the classical solution (pCCD) with the quantum solution (UpCCD via VQE). In Figure 5.11 we see such a comparison for the

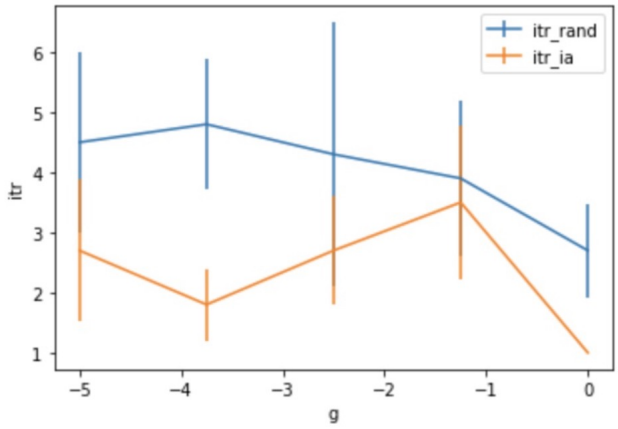


Figure 5.9: The number of iterations required to minimize the correlation energy (averaged over 10 trials) are compared for the pairing model with $P = 4$ energy levels and $N = 2$ pairs of particles. itr_rand and itr_ia are the number of iterations required to minimize the correlation energy for random initial parameters and initial parameters informed by MBPT (5.96), respectively.

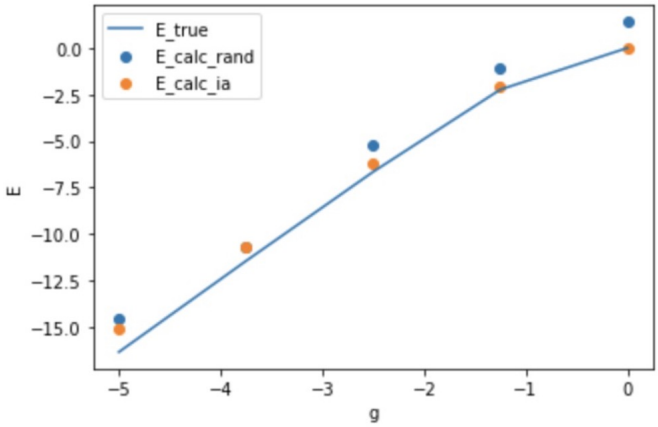


Figure 5.10: VQE calculated ground state correlation energies compared between the case of random initial parameterization (E_calc_rand) and the case of MBPT informed (5.96) initial parameterization (E_calc_ia).

pairing model with $P = 4$ energy levels, $N = 2$ pairs, linearly increasing singular particle energies $d_p = p$, and constant pairing strength $g_{pq} = g$. The VQE results were computed via a noiseless simulation on a classical computer. We see that for small values of g the

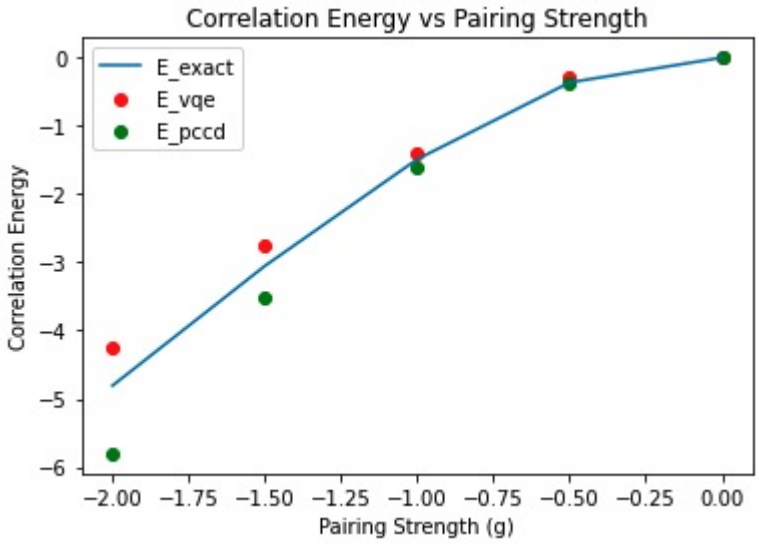


Figure 5.11: Pairing strength vs ground state correlation energy obtained through exact diagonalization (blue line), the variational quantum eigensolver (red dots) and pair coupled cluster doubles theory (green dots) for the pairing model with $P = 4$ energy levels, $N = 2$ pairs, linearly increasing singular particle energies $d_p = p$, and constant pairing strength $g_a^i = g$.

two methods perform relatively similarly. However, as the magnitude of g grows, we can clearly see that the quantum solution gives a better estimate of the ground state correlation energy than the classical solution.

Additionally, we test our VQE ansatz on a non-constant pairing strength pairing model. Here, we choose the separable function $g_a^i = g(ia/(\max(i)\max(a)))$ which is a more realistic approximation to real-world nuclei whose nuclear forces are often modeled to be separable. The function is divided by the maximums of the energy level indices i and j so that the values of g don't vary too wildly. In Figure 5.12 we see how pCCD performs well for small absolute values of g but fails for larger values, unlike VQE which gets close to the true correlation energy values for all values of g.

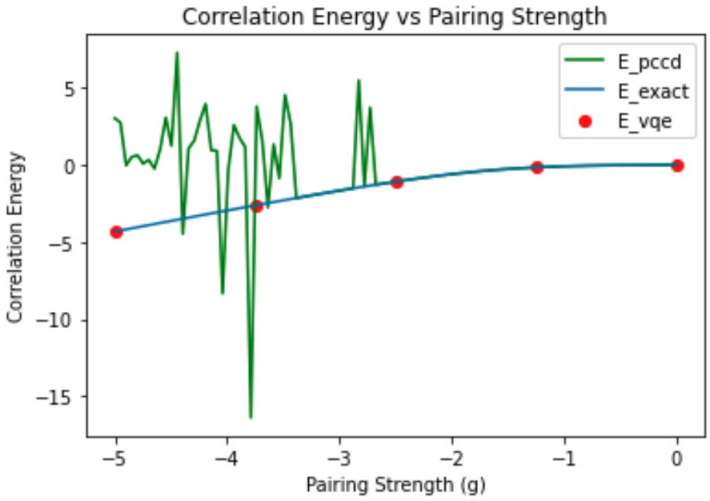


Figure 5.12: Pairing strength vs ground state correlation energy obtained through exact diagonalization (blue line), the variational quantum eigensolver (red dots) and pair coupled cluster doubles (green dots) for the pairing model with $P = 4$ energy levels, $N = 2$ pairs, linearly increasing singular particle energies $d_p = p$, and constant separable pairing strength $g_a^i = g(ia/(\max(i)(\max(a))$.

5.3.6 Iterative Quantum Excited States Algorithm

The following is a novel method to compute the eigenstates of a Hamiltonian and their corresponding energies. Let $|\psi_k\rangle$ represent the k^{th} eigenstate of the Hamiltonian H with energy E_n. That is $H|\psi_n\rangle = E_n|\psi_n\rangle$. Assume that $|\psi_k\rangle$ can be parameterized by the set of angles θ_k, that is $|\psi_k\rangle = |\psi(\theta_k)\rangle$. This algorithm is iterative, meaning that to find E_k, one must first use the algorithm to compute E_0 through E_k and θ_0 through θ_k. Assuming that one has already done this, recall that any state $|\psi(\theta)\rangle$ can be expanded in terms of the N eigenstates of H. That is

$$|\psi(\theta)\rangle = \sum_{n=0}^{N} \langle\psi(\theta_n)|\psi(\theta)\rangle\,|\psi(\theta_n)\rangle\,. \tag{5.100}$$

Then, note that the expectation value of H of an arbitrary state $|\psi(\theta)\rangle$ can be written as

$$\langle\psi(\theta)|H|\psi(\theta)\rangle \tag{5.101}$$

$$= \sum_{n,m=0}^{N} \langle\psi(\theta)|\psi(\theta_m)\rangle \langle\psi(\theta_n)|\psi(\theta)\rangle \langle\psi(\theta_m)| H |\psi(\theta_n)\rangle \tag{5.102}$$

$$= \sum_{n,m=0}^{N} E_n \langle\psi(\theta)|\psi(\theta_m)\rangle \langle\psi(\theta_n)|\psi(\theta)\rangle \langle\psi(\theta_m)|\psi(\theta_n)\rangle \tag{5.103}$$

$$= \sum_{n,m=0}^{N} E_n \langle\psi(\theta)|\psi(\theta_m)\rangle \langle\psi(\theta_n)|\psi(\theta)\rangle \delta_{nm} \tag{5.104}$$

$$= \sum_{n=0}^{N} E_n |\langle\psi(\theta_n)|\psi(\theta)\rangle|^2, \tag{5.105}$$

which allows

$$\langle\psi(\theta_k)|H|\psi(\theta_k)\rangle - \sum_{n=0}^{k-1} E_n |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 \tag{5.106}$$

$$= \sum_{n=0}^{N} E_n |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 - \sum_{n=0}^{k-1} E_n |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 \tag{5.107}$$

$$= \sum_{n=k}^{N} E_n |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 \tag{5.108}$$

$$\geq E_k \sum_{n=k}^{N} |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 \tag{5.109}$$

$$= E_k \left[1 - \sum_{n=0}^{k-1} |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 \right], \tag{5.110}$$

since $\sum_{n=0}^{N} |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2 = 1$. This implies that

$$E_k = \min_{\theta_k} \frac{\langle\psi(\theta_k)|H|\psi(\theta_k)\rangle - \sum_{n=0}^{k-1} E_n |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2}{1 - \sum_{n=0}^{k-1} |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2}. \tag{5.111}$$

The denominator can cause one trouble as small variations near zero would cause wild jumps in E_k. To get around this, we can instead minimize the following function

$$E_k = \min_{\theta_k} \langle\psi(\theta_k)|(H - \Delta)|\psi(\theta_k)\rangle - \sum_{n=0}^{k-1} E_n |\langle\psi(\theta_n)|\psi(\theta_k)\rangle|^2, \tag{5.112}$$

where here Δ is a large positive number which hopefully shifts the entire energy spectrum of H negative. Note that to apply this algorithm, one must be able to calculate the following two types of quantities:

- The expectation value of H in the i^{th} state $|\psi(\theta_i)\rangle$:

$$\langle\psi(\theta_i)|H|\psi(\theta_i)\rangle\,, \tag{5.113}$$

- The absolute square of the overlap between two states $|\psi(\theta_i)\rangle$ and $|\psi(\theta_j)\rangle$:

$$\left|\left\langle\psi(\theta_i)\middle|\psi(\theta_j)\right\rangle\right|^2. \tag{5.114}$$

The process to calculate the first quantity, the expectation value of H, is explained in section 3.4.2. The technique to efficiently calculate the second quantity, the absolute square of the overlap between two states, is explained in Appendix E.

As the algorithm is iterative, one starts with the case $k = 0$, in which case the algorithm reduces to the variational quantum eigensolver for the ground state energy

$$\min_{\theta_0} \langle\psi(\theta_0)|H|\psi(\theta_0)\rangle = E_0. \tag{5.115}$$

After obtaining θ_0 and E_0 through minimization, one uses these quantities to apply the algorithm for $k = 1$.

$$\min_{\theta_1} \langle\psi(\theta_1)|(H-\Delta)|\psi(\theta_k)\rangle - E_0|\langle\psi(\theta_0)|\psi(\theta_0)\rangle|^2 = E_1. \tag{5.116}$$

After minimizing, one obtains θ_1 and E_1. One then continues this process for $k = 2, ..., N$, ultimately resulting in knowledge of the parameters $\theta_0, ..., \theta_N$ that paramaterize all of the eigenstates of H and all of their corresponding energies $E_0, ..., E_N$. Thus the eigenspectrum

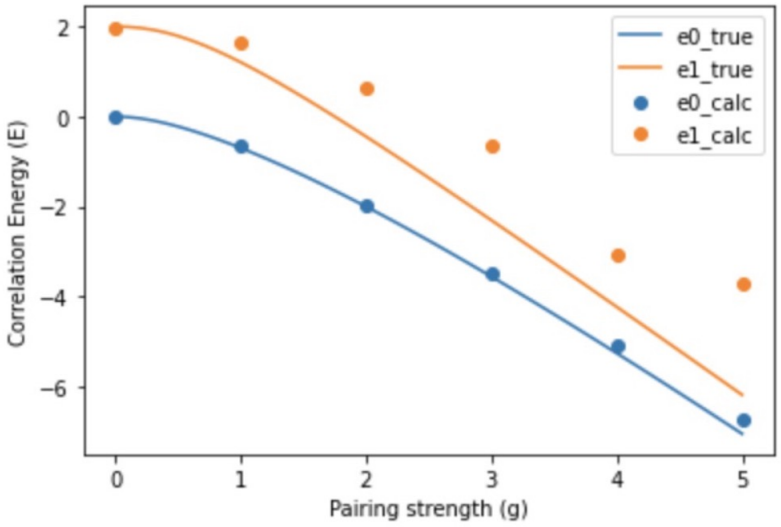


Figure 5.13: Comparison of correlation energies calculated using the iterative quantum excited states algorithm and direct diagonalization for the pairing model with $P = 4$ energy levels, $N = 2$ pairs, linearly increasing singular particle energy $d_p = p$, and constant pairing strength $g_a^i = g$.

of H can be fully calculated. We have tested this algorithm for the first two states of the pairing model with $(P, N) = (4, 2)$. In Figure 5.13, one can see that while the algorithm estimates the ground state quite well, the estimation for the first excited states grows worse for larger values of g. We hypothesize that this is due to the accumulation of errors present in this algorithm (5.112) for large values of g as such values make it harder for the minimization algorithm to succeed. If the algorithm's estimate for the ground state energy and/or the ground state are slightly off, this error will become compounded in estimations for the energies of higher energy levels.

5.3.7 Multi-Configuration Method

We have observed that in the constant g case, classical pCCD (and to a lesser but still significant extent) VQE with the UpCCD ansatz, have a harder and harder time finding

the true ground state as g grows. This is because the eigenfunction of the ground state becomes more and more entangled as g increases and begins to overtake the strength of the single-particle energies d_p. Such entangled states heuristically take a larger circuit depth to employ. In the limiting case of $g \to \infty$, the one-body term of H can be disregarded, leading to a Hamiltonian that mixes states and yet is not state dependent (g is constant). Heuristically, such a Hamiltonian would have, as its ground state, the equal superposition of all possible states, due to symmetry. Such a state is called a Dicke state. Recall that the Dicke state [23] has the following recursive form

$$\left|D_k^n\right\rangle = \sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + \sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle\,. \tag{5.117}$$

We will now prove that the Dicke state $\left|D_k^n\right\rangle$ is indeed an eigenvector of V (the limiting case of H where $g \to \infty$) where

$$V_n = g\sum_{p\neq q=1}^{n} A_p^\dagger A_q. \tag{5.118}$$

Specifically

$$V_n\left|D_k^n\right\rangle = gk(n-k)\left|D_k^n\right\rangle. \tag{5.119}$$

The proof is by induction. First, we prove the base case $(n,k) = (2,1)$.

$$V_2\left|D_1^2\right\rangle = \frac{g}{\sqrt{2}}(A_0^\dagger A_1 + A_1^\dagger A_0)(|10\rangle + |01\rangle) = g\left|D_1^2\right\rangle. \tag{5.120}$$

Next, we assume the induction hypothesis

$$V_n\left|D_k^{n-1}\right\rangle = gk(n-k-1)\left|D_k^{n-1}\right\rangle, \tag{5.121}$$

and use it to prove the statement as follows

$$V_n\left|D_k^n\right\rangle = g\left(\sum_{p\neq q=1}^{n} A_p^\dagger A_q\right)\left(\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + \sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle\right) \tag{5.122}$$

$$= g\left[\sum_{p\neq q=1}^{n-1} A_p^\dagger A_q + A_n^\dagger\left(\sum_{q=1}^{n-1} A_q\right) + \left(\sum_{p=1}^{n-1} A_p^\dagger\right) A_n\right]$$
$$\times\left[\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + \sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle\right] \tag{5.123}$$
$$= g\left(\sum_{p\neq q=1}^{n-1} A_p^\dagger A_q\right)\left(\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + \sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle\right)$$
$$+ g\left[A_n^\dagger\left(\sum_{q=1}^{n-1} A_q\right)\sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle + \left(\sum_{p=1}^{n-1} A_p^\dagger\right) A_n\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle\right] \tag{5.124}$$
$$= (k-1)(n-k)\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + k(n-k-1)\sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle$$
$$+ (n-k)\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + k\sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle \tag{5.125}$$
$$= k(n-k)\left(\sqrt{\frac{k}{n}}\left|D_{k-1}^{n-1}\right\rangle|1\rangle + \sqrt{\frac{n-k}{n}}\left|D_k^{n-1}\right\rangle|0\rangle\right) \tag{5.126}$$
$$= k(n-k)\left|D_k^n\right\rangle. \tag{5.127}$$

In this proof we have used the following lemmas:

$$\sum_{p=1}^{n} A_p^\dagger\left|D_k^n\right\rangle = \sqrt{(n-k)(k+1)}\left|D_{k+1}^n\right\rangle, \tag{5.128}$$
$$\sum_{p=1}^{n} A\left|D_k^n\right\rangle = \sqrt{k(n-k+1)}\left|D_{k-1}^n\right\rangle, \tag{5.129}$$

which themselves are proven inductively in Appendix D.

We can see how close the actual ground state of the pairing model (with $d_p = p$) is to the Dicke state as g grows by plotting their overlap squared. In Figure 5.14) we see such a plot for various energy levels p (with $n = \lfloor p/2 \rfloor$). The plot shows that by the time g reaches about 2.5, the overlap squared of the two states is above 0.9 for all values of p

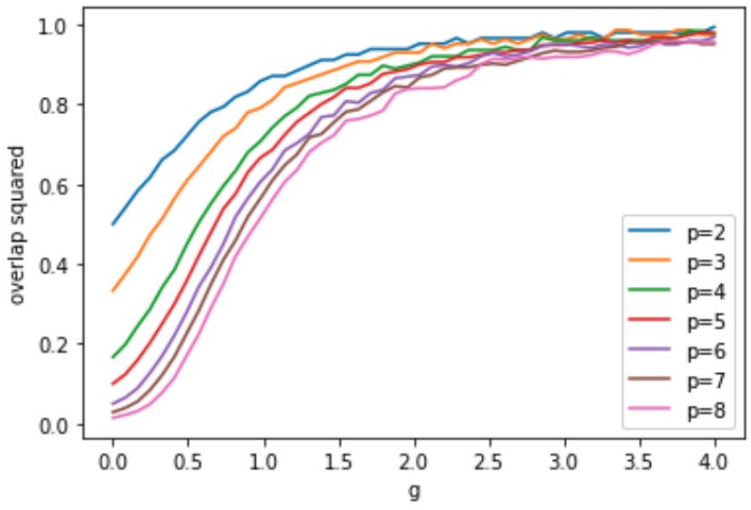
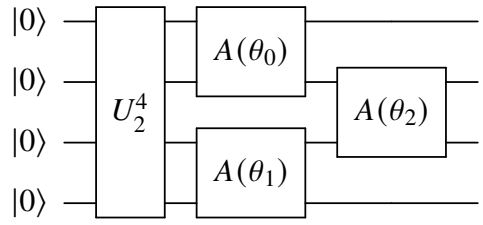


Figure 5.14: Plot of the overlap squared between the actual ground state and corresponding Dicke state for pairing models with various values of p and $n = \lfloor p/2 \rfloor$ over increasing values of constant pairing strength.

considered. It is also clear that the ground states for smaller values of p become closer to the Dicke state $\left|D_n^p\right\rangle$ at a faster rate than for larger values of p.

This suggests that, for sufficiently large g, one may consider the following ansatz for the pairing model: Use the deterministic [9] or variational (chapter 8) method to initialize one's quantum circuit to the Dicke state $\left|D_n^p\right\rangle$ for a pairing model with p energy levels and n pairs of fermions. Then, slowly increase the number of layers of variational gates in a "brick wall" fashion (6.1), running VQE on the ansatz each time, until the result no longer substantially improves. These extra variational gates are there to slightly perturb the Dicke state to hopefully approximate the true ground state of the pairing model. This is a so-called "adaptive" ansatz as it can change throughout the process of VQE. We tested such an ansatz for the pairing model with $P = 4$ energy levels and $N = 2$ pairs of fermions,

whose circuit representation is given below



(5.130)

where the gate U_2^4 prepares the corresponding Dicke state:

$$U_2^4 \left|0000\right\rangle = \left|D_2^4\right\rangle, \tag{5.131}$$

and A is defined the same as the two-qubit operators for the UpCCD ansatz (5.94) except that the their parameters θ_i (for $i = 0, 1, 2$) are completely free. That is

$$A(\theta) = \exp\left\{i\theta(X_a Y_i - Y_a X_i)\right\}, \tag{5.132}$$

which, as analyzed previously, restricts the minimization algorithm's search to the correct Hamming weight subspace. The results of using such an ansatz for VQE can be seen in Figures 5.15 and 5.16. The first shows a comparison between the VQE estimated and exact correlation energies while the second shows the relative error between the two from the first plot. As expected, this multi-configuration ansatz described above works better for larger absolute values of g, because we start in a state (the Dicke state) that has a significant overlap with the ground state. We name the ansatz the multi-configuration ansatz as it is initialized to the superposition of multiple initial configurations of the fermion pairs (or, bit-string states).

In future work, it would be of interest to explore if one can get away with preparing a Dicke like state (equal superposition of bit-string states but with incorrect phases, such as those prepared in chapter 8) or even partial Dicke states (the equal superposition of some of

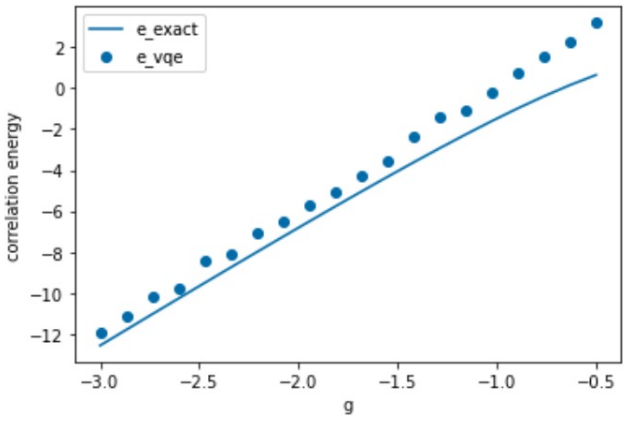


Figure 5.15: Plot of VQE estimated versus exact correlation energies for the pairing model with $P = 4$ energy levels and $N = 2$ pairs of fermions using the multi-configuration ansatz.

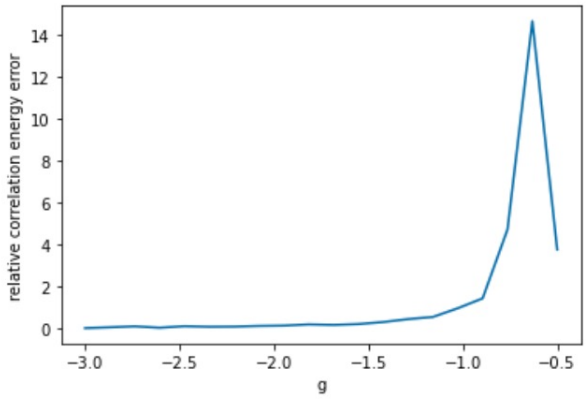


Figure 5.16: Plot of relative error between VQE estimated and exact correlation energies for the pairing model with $P = 4$ energy levels and $N = 2$ pairs of fermions using the multi-configuration ansatz.

the bit-string states that make up an entire Dicke states). One could also explore this ansatz for a larger number of energy levels P and different formulas of the single particle energies d_p. Finally, on might consider using the first few layers of the UpCCD ansatz to try to minimize the pairing model with $d_p = 0$ which should get one close to a Dicke state. Then add in the rest of the layers and minimize to the pairing model that one desires, initializing the initial layers to the parameters that the minimization algorithm found approximated the Dicke state.

5.4 Conclusion

In this section, we introduced the pairing model, a toy model for many-body nuclear physics upon which many new techniques are tested. First, we explored ways to solve the problem on a classical computer, including through the use of many-body perturbation theory and coupled cluster theory. These two methods served to help us in the next section by informing a good set of initial guesses for our variational circuit. In the aforementioned section, we walked through how to map the pairing model Hamiltonian and an extension of the unitary coupled cluster ansatz from pair fermionic operators to spin operators via an extension to the Jordan-Wigner transformation. We then presented our results of applying VQE to the pairing model and bench-marked it against classical coupled cluster theory. We went on to extend VQE in two novel ways: first, we introduced the novel iterative quantum excited states algorithm to search for the energy levels of excited states of the pairing model. We then introduced the novel multi-configurational ansatz for pairing models with large constant pairing strengths g. In the future, one might like to compare our excited states algorithms to similar algorithms and also test various relaxed forms of the multi-configurational ansatz that involve the initialization of Dicke-like states. Finally, one may

wish to apply the VQE solution to the pairing problem to approximate an actual nucleus that has strong pairing interactions such as ones with doubly-magic shells with valence electrons of a single type. Here, we have given a springboard off of which future research may be accomplished as quantum computing tackles ever more complex many-body nuclear physics systems.

CHAPTER 6

COLLECTIVE NEUTRINO OSCILLATIONS

6.1 Introduction

The next system we will consider is that of collective neutrino oscillations, which has Hamiltonian that is mathematically similar to that of the previous pairing models considered. Instead of using quantum algorithms to determine the energy spectrum of the Hamiltonian, as done in the two previous cases, we will instead use them to simulate the time-evolution of the system and measure its entanglement properties. Neutrinos are nuclear particles in the sense that they interact via the weak nuclear force (and gravity). The motivation for studying this system comes from the flavor evolution of neutrinos in dense astrophysical environments. It has been pointed out by Pantelone, Raffelt, and Sigl [61, 76] and others that neutrinos can exchange their flavors through forward scattering. If one starts with an anisotropic initial distribution of energy and/or angle (as found in supernovae, neutron star mergers, or the early universe), then the neutrino energy flux versus energy and flavor may be impacted by this non-trivial quantum many-body evolution. This can in turn affect the dynamics of these environments and other observables, including nucleosynthesis in the ejected material [25, 18].

Most often, these quantum equations have been treated on the mean-field level by replacing one of the spin operators in equation (6.1) by its expectation value, yielding a set of non-linear coupled differential equations. This makes the calculations tractable for several hundred energies and angles on modern computers [26]. More recently, studies of neutrino propagation as a quantum many-body problem have appeared, including for example [10, 35, 72, 62, 71, 16, 68, 67]. These works highlight the importance of

understanding the role of quantum correlations, such as entanglement, in order to quantify beyond mean-field effects in out-of-equilibrium neutrino simulations. A direct solution of the Schrödinger equation in equation (2.13), for a system of N configurations in energy and angle, incurs a computational cost that is exponential in N. This has limited early explorations of the problem to systems with $N = O(10)$ neutrinos. An alternative to reach larger system sizes, explored recently [68, 67], employs a matrix product state representation for $|\Phi(t)\rangle$ which allows one to track the exact time evolution in situations where entanglement never grows too much. For conditions leading to strong entanglement instead, simulations on digital/analog quantum computers have the potential to tackle the full neutrino dynamics while still enjoying a polynomial computational cost in system size N [54].

In this section, we explore the time-dependent many-body evolution and entanglement of neutrinos on a current-generation digital quantum computer. In section 6.2 we introduce in more detail the SU(2) spin model used to describe collective neutrino oscillations and an implementation of the time evolution operator appearing in equation (2.13) suitable for an array of qubits with linear connectivity. We present the results obtained for a small system with $N = 4$ neutrino amplitudes in section 6.4 and provide a summary and conclusions in section 6.8.

6.2 Hamiltonian

The Hamiltonian for neutrino flavor evolution in a dense neutrino environment includes three terms:

1. H_v | the vacuum mixing that has been determined from solar and accelerator neutrino experiments [38]

2. H_s | the forward scattering in matter leading to the well known MSW effect [87, 57]

3. H_n | neutrino-neutrino forward scattering

The first simplification we make is to truncate the neutrino flavors involved from three to two. That is, we only consider the oscillation of two flavors which, without loss of generality, we choose to be ν_e and ν_x. Here ν_x describes either ν_μ or ν_τ which we assume evolve similarly. This simplification allows for the neutrino-neutrino interaction H_n to be proportional to the dot product $\sigma_i \cdot \sigma_j$ of the SU(2) Pauli matrices describing the different flavor amplitudes of the two neutrinos

$$(H_n)_{ij} \propto \left(1 - \frac{q_i \cdot q_j}{\|q_i\|\|q_j\|}\right) \sigma_i \cdot \sigma_j \ . \tag{6.1}$$

where q_k is the momentum of the k-th neutrino and $\sigma_k = (X_k, Y_k, Z_k)$ is the vector of Pauli operators acting on its amplitude. The neutrino-neutrino interaction can exchange flavors of two neutrinos and, as can be seen above, has a forward scattering amplitude that depends on the angle between their momenta. Generalization to the three-flavor case is straightforward in principle.

We are working in the neutrino flavor basis, whose fermionic operators are related to those of the mass basis via the rotation

$$\begin{pmatrix} a_e(p) \\ a_x(p) \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a_1(p) \\ a_2(p) \end{pmatrix}. \tag{6.2}$$

In this basis the vacuum term H_v includes diagonal contributions describing the mass differences between different neutrino flavors and an off-diagonal term characterized by a mixing angle θ_v, while forward scattering term H_s is diagonal in the flavor basis.

For the simplified two-flavor case studied here, the state of the system can be described as an amplitude for a neutrino of each energy E_i (equal to the magnitude of momentum $\|q_i\|$) and direction of momentum (denoted by q_i), with $\alpha_\uparrow$ and $\alpha_\downarrow$ describing the amplitude of being in the electron flavor or in the heavy x (μ or τ) flavor respectively. These two amplitudes can be encoded in an SU(2) spinor basis. In this basis, the Hamiltonian can be written in terms of Pauli operators as the sum of a one-body term, describing both vacuum oscillations and forward scattering in matter,

$$H_1 = \frac{1}{2}\sum_i \left[(-\Delta_i \cos 2\theta_v + A)\,\sigma_i^z + \Delta_i \sin 2\theta_v \sigma_i^x\right], \tag{6.3}$$

and a two-body term, coming from the neutrino-neutrino forward-scattering potential V_{ij} from equation (6.1), which takes the following form [62]

$$H_2 = \sum_{i<j} \eta[1 - \hat{q}_i \cdot \hat{q}_j]\sigma_i \cdot \sigma_j. \tag{6.4}$$

In the one-body term, θ_v represents the vacuum mixing angle, while the strength is given by $\Delta_i = \delta m^2/(2E_i)$ with δm^2 the mass squared difference for neutrinos of different flavor. The matter potential enters as the diagonal contribution in the one-body term through the constant $A = \sqrt{2}G_F n_e$, with G_F the Fermi coupling constant and n_e the electron density.

As described in the introduction, the two-body term is a sum over spin-spin interactions with a coupling depending upon the relative angle between them. The overall strength depends on the neutrino density as

$$\eta = \frac{G_F}{\sqrt{2}V} = \frac{G_F n_\nu}{\sqrt{2}N}, \tag{6.5}$$

with N the number of neutrino momenta considered, given by the neutrino density n_ν times the quantization volume V. We can transform from fermionic flavor operators to flavor

isospin operators via the Jordan-Schwinger mapping:

$$J_p^+ = a_e^\dagger(p)a_x(p), \tag{6.6}$$

$$J_p^- = a_x^\dagger(p)a_e(p), \tag{6.7}$$

$$J_p^z = \frac{1}{2}\left(a_e^\dagger(p)a_e(p) - a_x^\dagger(p)a_x(p)\right), \tag{6.8}$$

which obey the SU(2) commutation relations

$$[J_p^+, J_q^-] = 2\delta_{pq}J_p^z \tag{6.9}$$

$$[J_p^z, J_q^\pm] = \pm\delta_{pq}J_p^{pm}, \tag{6.10}$$

and are thus isomorphic to the Pauli-spin matrices. Note that we can transform to Cartesian coordinates through the definition

$$J_p^\pm = J_p^x \pm iJ_p^y \tag{6.11}$$

The term of the Hamiltonian that describes vacuum oscillations H_ν is given by

$$H_\nu = \sum_p \left(\frac{m_1^2}{2p}a_1^\dagger(p)a_1(p) + \frac{m_2^2}{2p}a_2^\dagger(p)a_2(p)\right), \tag{6.12}$$

in the mass basis. To transform it to the Pauli basis, we first subtract it by the following term

$$\sum_p \frac{m_1^2 + m_2^2}{4p}\left(a_1^\dagger(p)a_1(p) + a_2^\dagger(p)a_2(p)\right), \tag{6.13}$$

which can be done without consequence as it is proportional to the identity (for a given number of particles) since the total number of neutrinos in each momentum mode is constant. This follows from the fact that forward scattering only exchanges the neutrino's momenta. Subtracting said term transforms the vacuum oscillation term (6.12) into

$$H_\nu = \sum_p \frac{\Delta m^2}{4p}\left(a_2^\dagger(p)a_2(p) - a_1^\dagger(p)a_1(p)\right), \tag{6.14}$$

where $\Delta_m^2 = m_2^2 - m_1^2$. Applying the inverse of the mapping between the flavor and mass bases (6.2) yields

$$H_\nu = \sum_p \frac{\Delta m^2}{4p}\Big[(\sin\theta a_e^\dagger(p) + \cos\theta a_x^\dagger(p))\,(\sin\theta a_e(p) + \cos\theta a_x(p)) \\ - (\cos\theta a_e^\dagger(p) - \sin\theta a_x^\dagger(p))\,(\cos\theta a_e(p) - \sin\theta a_x(p))\Big] \tag{6.15}$$

$$= \sum_p \frac{\Delta m^2}{4p}\Big[\sin 2\theta \left(a_e^\dagger(p)a_x(p) + a_x^\dagger(p)a_e(p)\right) \\ - \cos 2\theta \left(a_e^\dagger(p)a_e(p) - a_x^\dagger(p)a_x(p)\right)\Big], \tag{6.16}$$

which can be mapped to the Pauli basis via (6.6)-(6.8), resulting in

$$H_\nu = \sum_p B \cdot \sigma_p, \tag{6.17}$$

where

$$B = \frac{\Delta m^2}{2p}(\sin 2\theta, 0, -\cos 2\theta). \tag{6.18}$$

We've written the Hamiltonian in terms of Paul operators instead of the SU(2) operators $J^\pm$ and J^z as the two are isomorphic. The neutrino-neutrino scattering term is given by

$$H_{\nu\nu} = \frac{G_F}{\sqrt{2}V}\sum_{pq}(1 - \cos\phi_{pq})\big[a_e^\dagger(p)a_x(p)a_x^\dagger(q)a_e(q) + a_x^\dagger(p)a_e(p)a_e^\dagger(q)a_x(q) \\ + a_e^\dagger(p)a_e(p)a_e^\dagger(q)a_e(q) + a_x^\dagger(p)a_x(p)a_x^\dagger(q)a_x(q)\big], \tag{6.19}$$

which can also be mapped to the Pauli basis via (6.6)-(6.8), resulting in

$$H_{\nu\nu} = \frac{G_F}{\sqrt{2}V}\sum_{pq}(1 - \cos\phi_{pq})\left(\sigma_p^+\sigma_q^- + \sigma_q^-\sigma_p^+ + \sigma_p^z\sigma_q^z\right)$$

$$= \sum_{pq} J_{pq} \sigma_p \cdot \sigma_q, \tag{6.20}$$

with

$$J_{pq} = \frac{\sqrt{2} G_F}{V} (1 - \cos \phi_{pq}), \tag{6.21}$$

where we've disregarded the term $a_e^\dagger(p) a_e(p) a_x^\dagger(q) a_x(q) + a_x^\dagger(p) a_x(p) a_e^\dagger(q) a_e(q)$ as it is proportional to the identity. Note that, again, this is given in terms of Pauli spin matrices as opposed to angular momentum operators as the two are isomorphic as they obey the same SU(2) commutation relations. Putting the two terms together, the Hamiltonian for collective neutrino oscillations becomes

$$H = \sum_p B \cdot \sigma_p + \sum_{p<q} J_{pq} \sigma_p \cdot \sigma_q, \tag{6.22}$$

where we're able to restricted to sum to $p < q$ as the $p = q$ term is proportional to the identity and restricting $p \neq q$ to $p < q$ only picks up a factor of two which can be absorbed into the constant J_{pq}. We note that the Hamiltonian is similar to the Heisenberg model except that the two-body term is all-to-all rather than nearest neighbor; that is, it sums over both p and q rather than summing over p and $q = p + 1$. Its coupling strength $\eta \propto 1/N$ assures that the energy of the system is extensive. This allows us to obtain a well-defined many-body solution, in the limit of large numbers of neutrino momenta by extrapolating in system size N.

Because NISQ era quantum computers must have a relatively limited circuit depth in order to not be too noisy, the maximum time to which we can simulate time-evolution is also limited. Therefore, we consider a test case where the one-body and two-body interaction terms are set to be similar in magnitude, allowing flavor oscillations to occur

rapidly. An example of this case is the environment of order 100km from the surface of a proto-neutron star in a core collapse supernova. Here, the background matter density has decreased to the point where its contribution to the Hamiltonian is similar in magnitude to the neutrino-neutrino forward scattering. The relative angles of neutrino propagation are fairly small as neutrinos are emitted from a typical proto-neutron star radius of order 10km. In the neutrino bulb model [26] one further assumes that the evolution in a supernova depends only on the energy and the angle from the normal. Averaging over the azimuthal angles results in an average coupling $\langle 1 - \hat{q}_i \cdot \hat{q}_j \rangle = 1 - \cos(\theta_i)\cos(\theta_j)$.

For our test case, we take a monochromatic neutrino beam with energy $E_\nu = \delta m^2/(4\eta)$ and measure energies in units of the two-body coupling η. In order to avoid the symmetries introduced by the single angle approximation, we employ an anisotropic distribution of momentum directions using a simple grid of angles with

$$\phi_{pq} = \arccos(0.9)\frac{|p-q|}{N-1} \ . \tag{6.23}$$

This is similar to the standard bulb model as the relative couplings $1 - \cos\phi_{pq}$ are small. Additionally, we choose $\theta = 0.195$ so that the one-body and two-body terms are of relative strength. This leads the parameters to have the following numerical values:

$$B = \left(\sqrt{1 - 0.925^2}, 0, -0.925\right) \tag{6.24}$$

$$J_{pq} = \left[1 - \cos\left(\arccos 0.9\frac{|p-q|}{N-1}\right)\right] . \tag{6.25}$$

6.3 Connection to the Pairing Model

Before we go any further, we'd like to show here how collective neutrino oscillations can be viewed as a pairing model, thus justifying its use as an application of the pairing model.

This connection can be seen by writing the collective neutrino oscillation Hamiltonian in the mass basis. To do so, we introduce the mass isospin operators

$$K_p^+ = a_1^\dagger(p)a_2(p) \tag{6.26}$$

$$K_p^- = a_2^\dagger(p)a_1(p) \tag{6.27}$$

$$K_p^z = \frac{1}{2}\left(a_1^\dagger(p)a_1(p) - a_2^\dagger(p)a_2(p)\right), \tag{6.28}$$

which are analogous to the flavor isospin operators (6.6-6.8). This allows the vacuum oscillation term (6.14) to be readily identified as

$$\begin{aligned} H_\nu &= \sum_p \frac{\Delta m^2}{4p}\left(a_2^\dagger(p)a_2(p) - a_1^\dagger(p)a_1(p)\right) \\ &= \sum_p \omega_p K_p^z, \end{aligned} \tag{6.29}$$

where

$$\omega = -\frac{\Delta m^2}{2p}. \tag{6.30}$$

In order to deal with the neutrino neutrino interaction term (6.17), we must find the mapping between flavor isospin operators and mass isospin operators which follow from the mapping between flavor fermionic operators and mass fermionic operators (6.2):

$$\begin{aligned} J_p^+ &= a_e^\dagger(p)a_x(p) \\ &= [\cos\theta a_1^\dagger(p) + \sin\theta a_2^\dagger(p)][\cos\theta a_2(p) - \sin\theta a_1(p)] \\ &= \cos^2\theta a_1^\dagger(p)a_2(p) - \sin^2\theta a_2^\dagger(p)a_1(p) - \frac{1}{2}\sin 2\theta[a_1^\dagger(p)a_1(p) - a_2^\dagger(p)a_2(p)] \\ &= \cos^2\theta K_p^+ - \sin^2\theta K_p^- - \sin 2\theta K^z, \end{aligned} \tag{6.31}$$

$$J_p^- = a_e^\dagger(p)a_x(p)$$

$$
\begin{aligned}
&= [\cos\theta a_2^\dagger(p) - \sin\theta a_1^\dagger(p)][\cos\theta a_1(p) + \sin\theta a_2(p)] \\
&= \cos^2\theta a_2^\dagger(p)a_1(p) - \sin^2\theta a_1^\dagger(p)a_2(p) - \frac{1}{2}\sin 2\theta[a_1^\dagger(p)a_1(p) - a_2^\dagger(p)a_2(p)] \\
&= \cos^2\theta K_p^- - \sin^2\theta K_p^+ - \sin 2\theta K^z, \qquad (6.32)
\end{aligned}
$$

$$
\begin{aligned}
J_p^z &= \frac{1}{2}\left(a_e^\dagger(p)a_e(p) - a_x^\dagger(p)a_x(p)\right) \\
&= \frac{1}{2}\{[\cos\theta a_1^\dagger(p) + \sin\theta a_2^\dagger(p)]\,[\cos\theta a_1(p) + \sin\theta a_2(p)] \\
&\quad - [\cos\theta a_2^\dagger(p) - \sin\theta a_1^\dagger(p)]\,[\cos\theta a_2(p) - \sin\theta a_1(p)]\} \\
&= \frac{1}{2}\{\cos 2\theta[a_1^\dagger(p)a_1(p) - a_2^\dagger(p)a_2(p)] + \sin 2\theta[a_1^\dagger(p)a_2(p) - a_2^\dagger(p)a_1(p)]\} \\
&= \cos 2\theta K_p^z + \frac{1}{2}\sin 2\theta(K_p^+ + K_p^-), \qquad (6.33)
\end{aligned}
$$

which can be expressed succinctly in matrix form, for reference, as

$$
\begin{pmatrix} J_p^+ \\ J_p^- \\ J_p^z \end{pmatrix} = \begin{pmatrix} \cos^2\theta & -\sin^2\theta & -\sin 2\theta \\ -\sin^2\theta & \cos^2\theta & -\sin 2\theta \\ \frac{1}{2}\sin 2\theta & \frac{1}{2}\sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} K_p^+ \\ K_p^- \\ K_p^z \end{pmatrix}. \qquad (6.34)
$$

This implies that the dot product of the total flavor isospin operators is equal to the dot product of the total mass isospin operators:

$$
\begin{aligned}
J_p \cdot J_q &= \frac{1}{2}(J_p^+ J_q^- + J_p^- J_q^+) + J_p^z J_q^z \\
&= \frac{1}{2}\begin{pmatrix} J_p^+ & J_p^- & J_p^z \end{pmatrix} \begin{pmatrix} J_p^- \\ J_p^+ \\ 2J_p^z \end{pmatrix} \\
&= \frac{1}{2}\begin{pmatrix} K_p^+ & K_p^- & K_p^z \end{pmatrix} \begin{pmatrix} \cos^2\theta & -\sin^2\theta & \frac{1}{2}\sin 2\theta \\ -\sin^2\theta & \cos^2\theta & \frac{1}{2}\sin 2\theta \\ -\sin 2\theta & -\sin 2\theta & \cos 2\theta \end{pmatrix} \begin{pmatrix} \cos^2\theta & -\sin^2\theta & -\sin 2\theta \\ -\sin^2\theta & \cos^2\theta & -\sin 2\theta \\ \sin 2\theta & \sin 2\theta & 2\cos 2\theta \end{pmatrix} \begin{pmatrix} K_p^- \\ K_p^+ \\ K_p^z \end{pmatrix}
\end{aligned}
$$

$$= \frac{1}{2}\begin{pmatrix} K_p^+ & K_p^- & K_p^z \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} K_p^- \\ K_p^+ \\ K_p^z \end{pmatrix}$$

$$= \frac{1}{2}(K_p^+ K_q^- + K_p^- K_q^+) + K_p^z K_q^z \tag{6.35}$$

$$= K_p \cdot K_q. \tag{6.36}$$

Thus, the neutrino-neutrino interaction term (6.20) can be written in terms of total mass isospin operators as

$$H_{\nu\nu} = \sum_{pq} J_{pq} K_p \cdot K_q. \tag{6.37}$$

Putting the terms (6.29 - 6.37) together gives us the Hamiltonian in terms of mass isospin operators

$$H = \sum_p \omega_p K_p^z + \sum_{pq} J_{pq} K_p \cdot K_q. \tag{6.38}$$

Consider now the single-angle approximation, where we assume that neutrinos traveling in different directions undergo the same flavor evolution, implying that the term $\sum_{pq} \cos\phi_{pq} J_p \cdot J_q$ (and therefore also $\sum_{pq} \cos\phi_{pq} K_p \cdot K_q$) averages to zero, which (after recalling $J_{pq} = \mu(1 - \cos\phi_{pq})$) leads to the Hamiltonian

$$H = \sum_p \omega_p K_p^z + \mu \sum_{pq} K_p \cdot K_q, \tag{6.39}$$

which can be written as

$$H = \sum_p \omega_p K_p^z + 2\mu \sum_{p<q} (K_p^x \cdot K_q^x + K_p^y \cdot K_q^y), \tag{6.40}$$

where we've dropped the terms $\mu \sum_{pq} K_p^z \cdot K_q^z$ and $\mu \sum_{p=q} (K_p^x \cdot K_q^x + K_p^y \cdot K_q^y) = 2\mu \sum_p I_p$ as they are both proportional to the identity (for a fixed number of neutrinos) and therefore have no effect on the time-evolution of the system. Recall now the pairing Hamiltonian in terms of Pauli spin operators (5.86) which in the case of constant pairing strength ($g_q^p = g$) we can write as

$$H = -\frac{1}{2} \sum_{p=1}^{P} (2d_p + g) Z_p + \frac{1}{2} g \sum_{\substack{p,q=1 \\ p<q}} (X_p X_q + Y_p Y_q), \tag{6.41}$$

by dropping the term $\frac{1}{2} \sum_{p=1}^{P} (2d_p + g) I_p$ as it is proportional to the identity. With this, we can see that, since the mass isospin operators and Pauli-spin operators are isomorphic (share the same commutation relations), the collective neutrino oscillation Hamiltonian in the single-angle approximation (6.40) is equivalent (up to terms proportional to the identity) to the pairing Hamiltonian with constant pairing strength (6.41) if we set $d_p = \omega_p - g/2$ and $g = 4\mu$. This implies that the two systems evolve identically in time.

6.4 Time Evolution

One of the challenges in implementing the time evolution of the collective neutrino oscillation Hamiltonian (6.22) on a quantum computer is to find an accurate approximation to the time evolution operator (2.14)

$$U(t) = \exp\{-iHt\}, \tag{6.42}$$

that can be decomposed efficiently into quantum gates [54]. In this work, we accomplish this by using a first-order Trotter-Suzuki decomposition [78] of the time evolution operator. This decomposition requires partitioning the Hamiltonian. A naive partitioning would be to simply keep the Hamiltonian (6.22) written as is which would lead to the following

approximation for the time-evolution operator

$$U_1(t) = \prod_{p=1}^{N} e^{-itB\cdot\sigma_p} \prod_{p<q=1}^{N} e^{-itJ_{pq}\sigma_p\cdot\sigma_q}. \tag{6.43}$$

Instead, we choose the pair propagation partition, in which we rewrite the Hamiltonian (6.22) as a strictly two-body term

$$H = \sum_{p<q}^{N} h_{pq}, \tag{6.44}$$

where

$$h_{pq} = \frac{1}{N-1} B \cdot (\sigma_p + \sigma_q) + J_{pq}\sigma_p \cdot \sigma_q, \tag{6.45}$$

which is permitted, as

$$\begin{aligned}
\sum_{p<q}^{n}(\sigma_p + \sigma_q) &= \sum_{p=1}^{n}\sum_{q=p+1}^{n}(\sigma_p + \sigma_q) \\
&= \sum_{p=1}^{n}\sum_{q=p+1}^{n}\sigma_p + \sum_{q=2}^{n}\sum_{p=1}^{q-1}\sigma_q \\
&= \sum_{p=1}^{n}(n-p)\sigma_p + \sum_{q=1}^{n}(q-1)\sigma_q \\
&= \sum_{p=1}^{n}\left[(n-p) + (p-1)\right]\sigma_p \\
&= (n-1)\sum_{p=1}^{n}\sigma_p.
\end{aligned} \tag{6.46}$$

This leads to the pair-propagation time-evolution operator approximation U_2, defined as

$$U_2(t) = \prod_{p<q}^{N} u_{pq} \tag{6.47}$$

where

$$u_{pq} = e^{-ith_{pq}}, \tag{6.48}$$

which is correct up to additive error $\epsilon = \mathcal{O}(t^2)$. This partitioning is motivated by past experience with the Euclidean version of this evolution operator in quantum Monte Carlo which suggests that this partitioning yields a better approximation to the time-evolution operator $U(t)$ (see eg. [15, 14]). The main reason we choose this partitioning, however, is that it has better error scaling than the original partitioning U_1, as detailed below. While asymptotic scaling of the approximation error ϵ is quadratic in the time-step t for both approximations [78], the pair approximation is expected to perform better in practice for cases where an accurate description of pair evolution is important due, for instance, to strong cancellations between the one-body and two-body contributions in the Hamiltonian. In the neutrino case, these situations can occur with appropriate initial conditions so that, for typical states in the evolution, we have for most pairs that

$$\left|\langle K_{pq}\rangle + \langle V_{pq}\rangle\right| \ll \left|\langle K_{pq}\rangle\right| + \left|\langle V_{pq}\rangle\right| \tag{6.49}$$

where we have used the short-hand

$$\langle K_{pq}\rangle = \frac{1}{N-1} B \cdot \langle \sigma_p + \sigma_q \rangle \tag{6.50}$$

$$\langle V_{pq}\rangle = J_{pq} \langle \sigma_p \cdot \sigma_q \rangle. \tag{6.51}$$

Since the difference between the two approximations is not expected to hold for a generic initial state, standard error measures like the matrix norm of the difference with the exact propagator

$$\left\| \exp\left(-itH\right) - U_{1/2}(t) \right\| \tag{6.52}$$

are not expected to capture the effect. This is, in fact, found in practice for our system. In panel (a) of Figure 6.1 we show the estimate from equation (6.52) for the $N = 4$ neutrino model considered in this work. This error estimate indicates that the U_1 approximation has

a smaller maximum error than U_2 up to long times. We can look at a more direct measure of accuracy for our specific setup by considering instead the state fidelity

$$f(t) = \left|\langle \Psi_{1,2}(t)|\Psi(t)\rangle\right|^2 \tag{6.53}$$

between the exact state $|\Psi(t)\rangle$ at time t and one of its approximations $\left|\Psi_{1,2}(t)\right\rangle$, defined as

$$\left|\Psi_{1,2}(t)\right\rangle = U_{1,2}\,|\Psi(0)\rangle\,. \tag{6.54}$$

We show $f(t)$ for both approximations in panel (b) of Figure 6.1. The result here suggest that instead the pair approximation produces a state with a higher fidelity with the true state than the simple linear propagator U_1, especially at relatively long time-steps $t \in [4, 8]$.

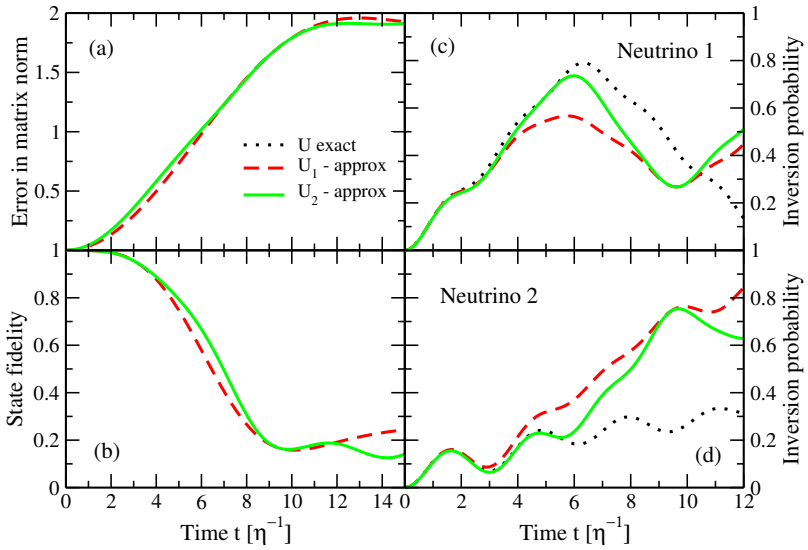


Figure 6.1: Panel (a) shows the error in matrix 2-norm equation (6.52) of the two approximations U_1 and U_2 described in the text. Panel (b) shows the state fidelity and the right panels show results for the inversion probability $P_{\rm inv}(t)$. Panel (c) is for neutrino one while panel (d) is for neutrino 2.

Finally, since we are mostly interested in flavor observables diagonal in the computational basis, we also show a direct comparison of the inversion probability for two out of

the $N = 4$ neutrinos using both approximations and the exact propagator (panels (c) and (d)). The details of how we compute the inversion probability are forthcoming but these selected results were included here first to show more clearly that the pair approximation allows us to correctly describe the evolution of flavor for substantially longer times than the U_1 approximation. The results reported here do depend on the specific choice of ordering of qubits in the time evolution layers shown in Figure 6.2. In both the present analysis and the simulation results we used the best ordering which we empirically found to be $(1, 3, 2, 4)$ as one would've expected based on the initial state and the criterion equation (6.49) above. A more rigorous discussion of the relative accuracy between the canonical first order and the pair approximation, together with the effect of ordering choices, could be explored in future work.

Because our Hamiltonian (6.44) sums over all $p < q$, a naive implementation would require either a device with all-to-all qubit-connectivity (like a trapped ion systems [58]) or an extensive use of the SWAP gate (3.37). Recall that the SWAP operation exchanges the states of the two qubits upon which it acts. It can therefore be used to bring the states of any pair of qubits next to one another in the physical qubit space, allowing one to apply two-qubit interactions between them on a device with linear qubit-connectivity. This would naively require a sequence of order N SWAP gates. However, we will show that, since we need to apply all possible pair interactions, it is actually possible to carry out a complete Trotter step of the time-evolution operator (6.47) without incurring any overhead due to the application of the SWAP operations. The scheme is inspired by the more general fermionic swap network construction presented in [51].

We illustrate this idea using the diagram shown in Figure 6.2 for a simple case with

$N = 4$ neutrinos. Starting from the initial state on the left, we first apply the unitaries u_{pq} from equation (6.47) to the odd bonds: for the $N = 4$ case, these are the bonds between the $(1, 2)$ and $(3, 4)$ pairs of qubits. Before moving to the next pairs, we also apply a SWAP operation to the same pairs upon which we just acted. The resulting unitary operation is denoted as a double line joining circles (qubits) in Figure 6.2 and the net effect is that at the next step the qubits that have interacted get interchanged. As any two-qubit unitary can be efficiently decomposed into a sequence of gates involving three CNOTs and 15 single qubit rotations (3.39), we see that by decomposing u_{pq}SWAP in this manner, we may apply the SWAP gates without incurring any additional gates or depth.

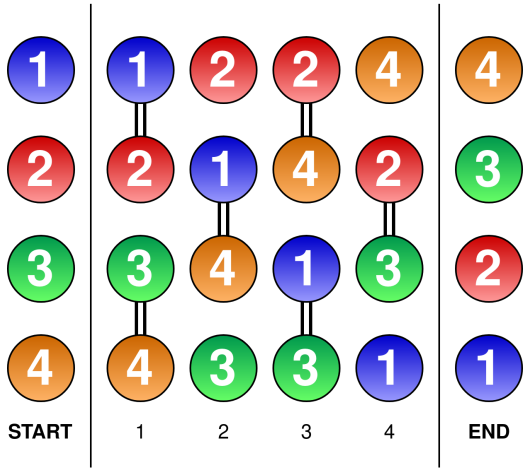


Figure 6.2: Pictorial representation of the swap network used in our simulation in the case of $N = 4$ neutrinos.

At the end of a sequence of N such combined operations we will have implemented the full unitary in equation (6.47) while, at the same, we inverted the ordering of qubits, as shown in Figure 6.2. This approach requires exactly the minimum number $\binom{N}{2}$ of nearest-neighbor pair operations, while the shifted ordering can be controlled completely, and efficiently, by classical means. Note that if we were to repeat at this point the same swap

network in reverse order, the full unitary will correspond to a second order step for time $2t$ and the final ordering of qubits will be restored to its original one. This is the strategy used in Refs. [68, 67] to study the neutrino Hamiltonian with matrix product states. In this first implementation on quantum hardware, we focus instead on a single, linear-order, time step.

Note that since we are only using nearest neighbor two-qubit gates, the total number of entangling gates required for a full time evolution step is bounded from above by $3\binom{N}{2}$ while the maximum number of single qubit operations is bounded by $15\binom{N}{2}$. As we will see in the results presented below, the presence of a large number of arbitrary single qubit rotations seems to be the limiting factor in implementing this scheme on the quantum device we used in this exploration.

In order to study the build up of correlations and entanglement generated by the time-evolution under the Hamiltonian in (6.22) we first initialize a system of $N = 4$ qubits to the following product state

$$|\Phi_0\rangle = |e\rangle \otimes |e\rangle \otimes |x\rangle \otimes |x\rangle = |0011\rangle \tag{6.55}$$

We then preform one step of time evolution for time t by applying the N layers of nearest-neighbor gates as described above. This corresponds to a single Trotter-Suzuki step for different values of the time-step t. The four SU(2) spins representing the neutrinos are mapped to qubits $(2, 1, 3, 4)$ on the IBMQ Vigo quantum processor [45], whose connectivity is schematically depicted in Figure 6.3. The resulting qubits are linearly connected, allowing us to natively carry out the complete simulation scheme depicted in Figure 6.2 above.

The first observable we compute is the inversion probability $P_{\text{inv}}(t)$, the probability that

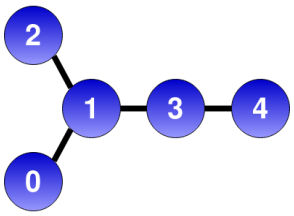


Figure 6.3: Layout of the IBM Quantum Canary Processor Vigo [45]. Shown are the five qubits, labeled from 0 to 4, and their connectivity denoted as solid black lines.

a neutrino is measured to not be in its original flavor state, of each individual neutrino as a function of time. Note that under the simultaneous exchanges $1 \leftrightarrow 4$ and $2 \leftrightarrow 3$, while the Hamiltonian in (6.22) is invariant, the flavor content of the initial state $|\Phi_0\rangle$ gets reversed. Therefore, in the limit of no error, $P_{\text{inv}}(t)$ should be the same for the pairs of neutrinos $(1,4)$ and $(2,3)$. The errors in the approximation of the propagator (6.47) do not exactly follow this symmetry, with deviations in the range $3-7\%$. We show the results for $P_{\text{inv}}(t)$ obtained via the approximate evolution operator $U_2(t)$ as solid black lines in Figure 6.4, for the pair $(1,4)$, and in Figure 6.5 for the pair $(2,3)$. The ideal, and symmetric, result is shown instead as a purple dashed line. We see that the approximation error is very small up to relatively large time $\eta t \approx 6$. As discussed earlier, this is in large part an effect of using the pair propagator $U_2(t)$ instead of the naive first order formula in equation 6.43.

The results shown in Figure 6.4 and Figure 6.5 were obtained using either the real quantum device (right panels denoted QPU) or a local virtual machine simulation employing the noise model implemented in Qiskit [3] (left panels denoted by VM) initialized with calibration data from the device. In both plots we report the results (denoted by [bare]) obtained directly from the simulation and including only statistical errors coming from a finite sample size (here and in the rest of the section we use 8192 repetitions, or "shots", for every data point), as well as results obtained after performing error mitigation (denoted

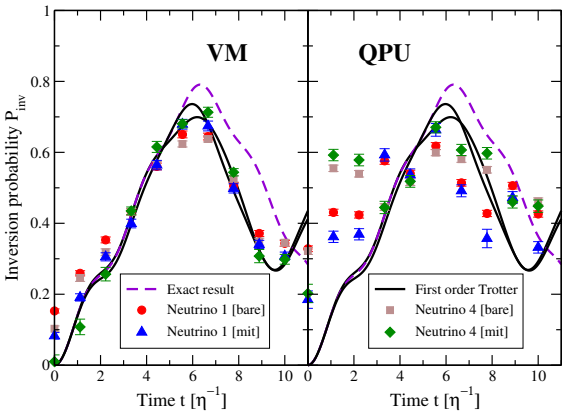


Figure 6.4: Inversion probability $P_{\mathrm{inv}(t)}$ for neutrinos one and four: the red circle and brown square correspond to the bare results, the blue triangle and the green diamond are obtained after error mitigation (see text). The left panel (VM) are virtual machine results while the right panel (QPU) are results obtained on the Vigo [45] quantum device.

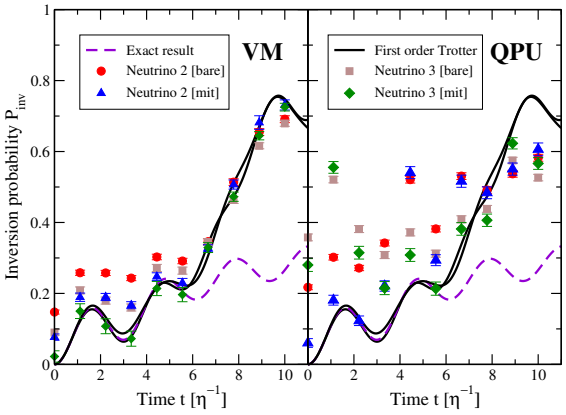


Figure 6.5: Inversion probability $P_{\mathrm{inv}}(t)$ for neutrinos two and three. The notation is the same as for Figure 6.4.

by [mit]). This corresponds to a final post-processing step that attempts to reduce the influence of the two main sources of errors: the read-out errors associated with the imperfect measurement apparatus and the gate error associated with the application of entangling gates. The latter error is dealt with using a zero noise extrapolation strategy (see [31, 29] and section 6.6 for additional details).

As seen in previous similar calculations [70, 69], the VM results obtained using the

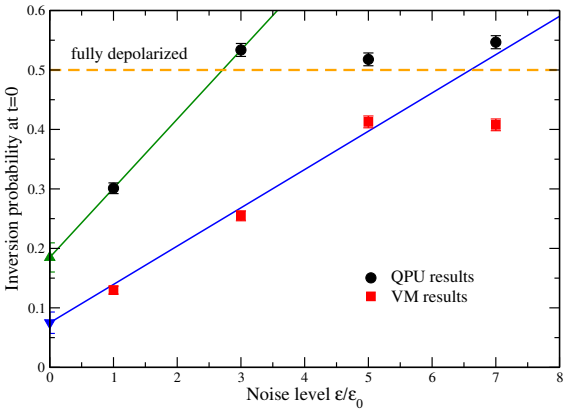


Figure 6.6: Inversion probability P_{inv} at the initial time $t = 0$ for the first neutrino. Black solid circles are results from the Vigo QPU [45] while the red squares correspond to results obtained using the VM with simulated noise. Also shown are extrapolations to the zero noise limit, for both the QPU (green line) and the VM (blue line), together with the extrapolated value (greed triangle up and blue triangle down respectively). The dashed orange line denotes the result for a maximally mixed state.

simulated noise are much closer to the ideal result than those obtained with the real device. This is also reflected in the fact that the error mitigation protocol is not as successful with the real QPU data as it is with the simulated VM data. This behavior is possibly linked to the substantial noise caused by the presence of a large number of single qubit operations (up to 90 degree rotations for time evolution and two for state preparation) together with the relatively large CNOT count of 18. In fact, the performance of error mitigation for the results with the largest state preparation circuits presented in [69] is superior to the one obtained here, despite the use of the same device, the same error mitigation strategy and a comparable number of entangling gates (15 CNOT in that case) while the number of rotations was only 14. This suggests coherent errors constitute a considerable fraction of the overall error seen in the results above.

In order to highlight the difficulties encountered when performing noise extrapolation for this data, we plot in Figure 6.6 the results obtained from both the QPU (black circles)

and the VM (red squares) for the inversion probability of the first neutrino at the initial time $t = 0$ together with a linear extrapolation using the first two points for the QPU (green line) and the first three points for the VM (blue line). The exact result is of course $P_{\text{inv}}(0) = 0$ and we see that neither strategy is able to predict the correct value. The horizontal dashed line is the value expected when the system is in the maximally mixed state, corresponding to full depolarization. As shown in the data, for the real QPU results, only the first level of noise extrapolation contains useful information and a more gentle noise amplification strategy, like the one proposed in [41], could provide a substantial advantage over the strategy adopted here.

6.5 Dynamics of entanglement

In order to track the evolution of entanglement in the system we perform complete state tomography for each of the six possible qubit pairs in our system by estimating, for each pair (k, q), the 16 expectation values

$$M^{k,q}_{\alpha,\beta}(t) = \langle \Phi(t) | P^{\alpha}_{k} \otimes P^{\beta}_{q} | \Phi(t) \rangle, \tag{6.56}$$

with $P^{\alpha}_{k} \in \{I, X, Y, Z\}$ being a Pauli matrix acting on the k^{th} qubit and $|\Phi(t)\rangle$ the state obtained from $|\Phi_0\rangle$ by applying the time-evolution operator as in equation (2.13). In principle, we may reconstruct the density matrix for the pair of qubits (k, q) directly from these expectation values as

$$\rho^{D}_{kq}(t) = \sum_{\alpha=1}^{4} \sum_{\beta=1}^{4} M^{k,q}_{\alpha,\beta}(t) P^{\alpha}_{k} \otimes P^{\beta}_{q} \ . \tag{6.57}$$

In practice however, we can only estimate the matrix elements $M^{k,q}_{\alpha,\beta}(t)$ to some finite additive precision, and the approximation in equation (6.57) is not guaranteed to be a physical density matrix (positive definite and trace equal to 1). In this work we use the

common approach (see eg. [6]) of performing a maximum-likelihood (ML) optimization, while enforcing the reconstructed density matrix $\rho^{ML}_{kq}(t)$ to be physical. We note in passing that it is possible to devise operator bases that are more robust than the choice used in equation (6.56) (see eg. [21]) but we don't explore this further in our work.

In order to propagate the effect of statistical errors into the final estimator for $\rho^{ML}_{kq}(t)$, we use a re-sampling strategy similar to what was introduced in [69] but using a Bayesian approach to determine the empirical posterior distribution. We provide a detailed description of the adopted protocol in subsection 6.6.1.

6.5.1 Entanglement Entropies

As we mentioned in the introduction, one of the main differences between a mean field description and the full many-body description of the dynamics of the neutrino cloud is the absence of quantum correlations, or entanglement, in the former. Past work on the subject [16, 71] looked at the single spin entanglement entropy defined as

$$S_k(t) = -\mathrm{Tr}\left[\rho_k(t)\log_2\left(\rho_k(t)\right)\right], \tag{6.58}$$

with $\rho_k(t)$ the reduced density matrix of the k-th spin. A value of the entropy $S_k(t)$ different from zero indicates the presence of entanglement between the k-th neutrino and the rest of the system.

In our setup, we compute the one-body reduced density matrix from the maximum-likelihood estimator of the pair density matrix defined above, explicitly

$$S^{ML}_{k;q}(t) = -\mathrm{Tr}\left[\rho^{ML}_{k;q}(t)\log_2\left(\rho^{ML}_{k;q}(t)\right)\right], \tag{6.59}$$

where the reduced density matrices are computed from

$$\rho^{ML}_{k;q}(t) = \mathrm{Tr}_q\left[\rho^{ML}_{kq}(t)\right], \tag{6.60}$$

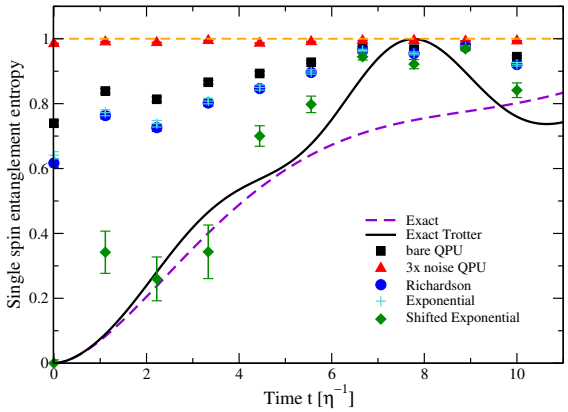


Figure 6.7: Single spin entanglement entropy for neutrino 2. Black squares are bare results obtained from the QPU, red triangles are results obtained by amplifying the noise to $\epsilon/\epsilon_0 = 3$, the blue circles are obtained using Richardson extrapolation, the turquoise plus symbols indicate results obtained using the standard exponential extrapolation and the green diamonds correspond to the results obtained from a shifted exponential extrapolation using the maximum value of the entropy (indicated as a dashed orange line).

and Tr_q denotes the trace over the states of the q-th qubit. We combine the three values obtained in this way for each neutrinos as follows: the estimator for the single-spin entanglement entropy is obtained from the average

$$S_k^{\mathrm{avg}}(t) = \frac{1}{3}\sum_q S_{k;q}^{ML}(t), \tag{6.61}$$

summing over pairs containing the k-th spin, while as an error estimate we use the average of the three errors.

As for the case of the inversion probability $P_{\mathrm{inv}}(t)$ studied in the previous section, the substantial noise present in the QPU data prevents us from using the full set of results at the four effective noise levels. In order to overcome this difficulty, we have performed zero noise extrapolations using only results for effective noise levels $r = \epsilon/\epsilon_0 = 1, 3$ and performed a Richardson extrapolation (in this case equivalent to a simple linear fit as done in [29]), a two point exponential extrapolation [31], and an exponential extrapolation

with shifted data. The latter technique consists in shifting the data for the entropy by -1 (its maximum value) so that the result, in the limit of large noise, tends to 0 instead of $\log_2(2) = 1$. We then shift back the result obtained after extrapolation. The exponential extrapolation method is well suited for situations where expectation values decay to zero as a function of the noise strength ϵ, while maintaining a consistent sign, and this shift allows us to make the data conform to this ideal situation (section 6.6 for more details on the method). The impact on the efficacy of the error mitigation is dramatic as can be seen in the results presented in Figure 6.7 for the entropy of the second neutrino (the entropies for the other neutrinos follow a similar pattern; see subsection 6.7 for all four results). The results with the standard exponential extrapolation are presented as the turquoise plus symbols; they are almost the same as those obtained using Richardson extrapolation (blue circles) and show a significant systematic error. On the contrary, the results obtained with the shifted exponential extrapolation (green diamonds) are much closer to the expected results with our pair propagator partition (solid black curve). We expect more general multi-exponential extrapolation schemes, like those proposed in Refs. [37, 13], to enjoy a similar efficiency boost in the large noise limit achieved with deep circuits.

Using the reconstructed pair density matrix $\rho^{ML}_{kq}(t)$, we can clearly also directly evaluate the entanglement entropy of the pair

$$S^{ML}_{kq}(t) = -\mathrm{Tr}\left[\rho^{ML}_{kq}(t)\log_2\left(\rho^{ML}_{kq}(t)\right)\right] \ . \tag{6.62}$$

In Figure 6.8 we show the result of this calculation for the pair $(1,2)$, which started as electron flavor at $t = 0$, and the pair $(2,4)$ which started instead as heavy flavor states.

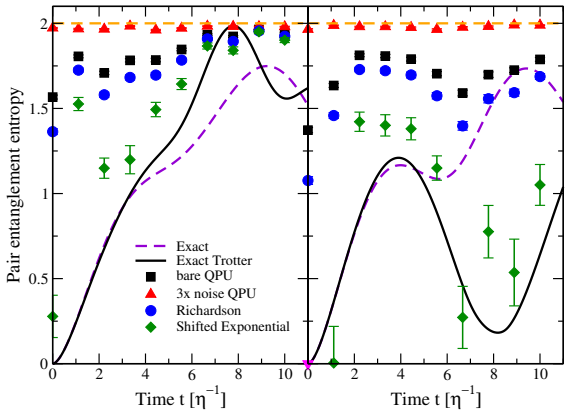


Figure 6.8: Pair entanglement entropy for the neutrino pair $(1,2)$ starting as $|e\rangle \otimes |e\rangle$ (left panel) and pair $(2,4)$ which starts as the flavor state $|e\rangle \otimes |x\rangle$ (right panel). Results obtained directly from the QPU are shown as black squares ($r = 1$) and red triangles ($r = 3$) while blue circles and green diamonds indicate mitigated results using Richardson and the shifted exponential extrapolations respectively. For the shifted exponential ansatz we use the maximum value of the entropy (indicated as a dashed orange line).The magenta triangle indicates a mitigated result with shifted exponential extrapolation below zero within errorbars.

6.5.2 Concurrence

In order to better understand these quantum correlations, we also compute the concurrence [88] for all the pair states. This measure of entanglement is defined for a two-qubit density matrix as

$$C(\rho) = \max\{0, \lambda_0 - \lambda_1 - \lambda_2 - \lambda_3\}, \tag{6.63}$$

where λ_i are the square roots of the eigenvalues, in decreasing order, of the non-Hermitian matrix

$$M = \rho\,(Y \otimes Y)\,\rho^*\,(Y \otimes Y), \tag{6.64}$$

with the star symbol indicating complex conjugation. The usefulness of this measure is its relation with the entanglement of formation [43, 88], which is the minimum number of maximally-entangled pairs needed to represent ρ with an ensemble of pure states [43].

The definition of concurrence in equation (6.63) does not lend itself as easily to be adapted in an error extrapolation procedure as the one we used to obtain the mitigated results in the previous sections. This is due to the presence of the max function in the definition of the concurrence: when the error is sufficiently strong to make the difference in eigenvalues

$$\widetilde{C}(\rho) = \lambda_0 - \lambda_1 - \lambda_2 - \lambda_3 \tag{6.65}$$

negative, the concurrence in equation (6.63) ceases to carry information about the error free result. For this reason, we will regard $\widetilde{C}$ as an "extended concurrence" which varies smoothly for large error levels and perform the truncation to positive values only after the zero noise extrapolation. The results obtained from the simulation on the Vigo QPU are shown in Figure 6.9 for two pairs of neutrinos: pair $(1, 2)$ starting as like spin at $t = 0$ and pair $(2, 4)$ which started as opposite flavors. The complete set of results for all pairs can be found in Figure 6.10 in subsection 6.7.

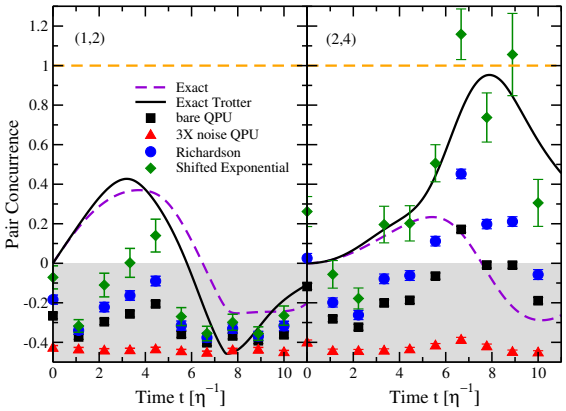


Figure 6.9: Extended concurrence $\widetilde{C}$ for two pairs of neutrinos, $(1, 2)$ in the left and $(2, 4)$ in the right panel. The convention for the curves and date point used here is the same as in Figure 6.8. The gray area indicates the region where the concurrence $C(\rho)$ is zero. The maximum value for the concurrence is shown as a dashed orange line.

The bare results are shown as black squares and we can immediately notice why the definition of $\widetilde{C}$ is so important in our case: the only bare data point with a measurable concurrence $C(\rho)$ is at $t \approx 6.7\eta^{-1}$ for pair $(2,4)$ (the right panel in Figure 6.9) while all the other results, including those obtained with a larger noise level (red triangles), are compatible with zero. In this situation, no mitigation of $C(\rho)$ would be possible.

By keeping the negative contributions, we see that the bare results often contain a substantial signal, while those at a higher error rate are already almost at the asymptotic value $\widetilde{C} = -0.5$ expected for a completely depolarized system. Note that our results seem to converge to a larger asymptotic value of $\widetilde{C} \approx -0.44$ instead of $\widetilde{C} = -0.5$. We can empirically explain this difference as the effect of statistical fluctuations. This allowed us to perform error extrapolation using both the Richardson and shifted exponential ansatz. Similarly to what we observed for the entanglement entropies in the previous section, the shifted exponential ansatz (with shift -0.5) produces consistently better results than Richardson extrapolation. This indicates that we are more close to the asymptotic large error regime than the small error limit used to motivate a polynomial expansion. The resilience of the exponential extrapolations to large errors, especially augmented by an appropriate shift, is seen here to be critical in extracting physical information from quantum simulations carried out near the coherence limit of the device used for implementation.

6.6 Error mitigation

6.6.1 Propagation of statistical uncertainties

In this section we describe the procedure we have adopted for propagating statistical errors in the results reported previously. We found that careful treatment of statistical errors was important for non linear functions of the expectation values like entropy and

concurrence of a reconstructed density matrix.

In the following, we will symbolically denote as $\langle O \rangle$, expectation values of Pauli operators which can be measured directly on the device. These are, for instance, the expectation values $\langle XX \rangle$, $\langle XY \rangle$, etc. needed to reconstruct a two-qubit density matrix.

We use a Bayesian approach to perform inference from the bare counts obtained from the device. The idea is best described initially for the simple case of a single qubit measurement. The probability of obtaining m measurements of the state $|1\rangle$ out of a total of M trials can be modelled as a binomial distribution

$$P_b(m; p) = \binom{M}{m} p^m (1-p)^{M-m}, \tag{6.66}$$

with p the probability of a $|1\rangle$ measurement. In order to infer the parameter p from a given sample m_i of measurement outcomes, we use Bayes' theorem

$$P(p|m_i) = \frac{P(m_i|p)P(p)}{\int dq P(m_i|q)P(q)} \ . \tag{6.67}$$

For the single qubit measurement, we use the binomial distribution as likelihood $P(m_i|p)$ and, in order to obtain a posterior $P(p|m_i)$ in closed form, we use the conjugate prior of the binomial: the beta distribution

$$P_\beta(p; \alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} p^{\alpha-1}(1-p)^{\beta-1} \ . \tag{6.68}$$

Here $\alpha, \beta > 0$ are the parameters defining the distribution and with $\alpha = \beta = 1$ we obtain a uniform distribution. The advantage of using the Beta distribution as a prior is that, after a measurement m_i of the system is available, the parameters (α_0, β_0) of the prior distribution get updated as

$$\alpha_i = \alpha_0 + m_i \quad \beta_i = \beta_0 + M - m_i \ . \tag{6.69}$$

Intuitively we can interpret the parameters (α_0, β_0) of the prior as assigning an a-priori number of measurements to the measurement outcomes, which are then updated as more measurements are performed. In this work we used a simple uniform prior corresponding to the choice $\alpha_0 = \beta_0 = 1$ for the prior parameters.

After the inference step described above, we calculate the expectation value of a generic non-linear function $\langle F[O] \rangle$ by sampling new outcomes m'_k using the posterior distribution. More in detail, we generate a new artificial measurement m'_k after the measured m_i by the following procedure

- sample a value p'_k from the posterior $P(p'_k|m_i)$
- sample a new measurement outcome m'_k from the likelihood $P_b(m'_k; p'_k)$

The new measurements m'_k obtained in this way are then samples from the predictive posterior distribution.

Using an ensemble of size L obtained in this way, we compute $\langle F[O] \rangle$ by taking an average of the results obtained for each individual sample

$$\langle F[O] \rangle \approx \frac{1}{L} \sum_{k=1}^{L} F[O_k] \ . \tag{6.70}$$

The error bars reported previously are 68% confidence intervals which we found in most cases where well approximated by a Gaussian approximation.

This scheme is complete only for single qubit measurements but a generalization to generic multi-qubit observables can be obtained in a straightforward way. In the situation where we are estimating expectation values over N qubits, the probability of measuring a specific collection of N bit strings m_i in M repeated trials can be described with a multinomial distribution with N probabilities. We use this distribution as the likelihood

$P(m_i|\vec{p})$ in Bayes theorem and, for similar reasons as above, we take its conjugate prior distribution: the Dirichlet distribution (also initialized as uniform as for the Beta above). The procedure we follow is otherwise exactly equivalent to what we described above.

6.6.2 Read-out mitigation

The qubit measurements on a real device are not perfect and it is therefore important to understand the associated systematic errors. We refer the reader to Appendix H.1 of [69] for a more detailed derivation of the exact procedure we employ and the motivations behind it. Here, we instead describe the main difference with the scheme described there which comes from the use of the Bayesian inference scheme described in the previous subsection.

In the calculations presented here, we work under the assumption that read-out errors are independent on each qubit and perform a set of $2N$ calibration measurements c_i (requiring two separate executions) to extract the parameters $(\vec{e}_0, \vec{e}_1)$ of the noise model (see Eq.(H1) of [69]). In order to consistently propagate the statistical uncertainties associated from the finite sample statistic used to estimate the noise parameters, we use an additional layer of Bayesian sampling using a binomial prior for the two error probabilities (e_0^n, e_1^n) associated to each qubit n.

Using a single pair of error probability vectors $\epsilon_i = (\vec{e}_0, \vec{e}_1)_i$, obtained either by direct measurement or by sampling from the posterior, we can generate a linear transformation $\mathbf{C}_i$ that maps a set of (in general multi-qubit) measurements m_i to a new set $\widetilde{m}_i$ with reduced read-out errors (see [69] for more details).

The complete procedure that we use to generate an ensemble of measurements $\{\widetilde{m}_i'\}$ with read-out mitigation starting from a single calibration measurement c_i and Pauli operator measurement m_i is as follows

- sample a value p'_k from the posterior $P(p'_k|m_i)$
- sample a new measurement outcome m'_k from the likelihood $P_b(m'_i; p'_k)$
- for each qubit $n = \{1, \ldots, N\}$
  - sample a pair (e'^n_0, e'^n_1) of error probabilities from the posterior $P(e^n_0, e^n_1|c_i)$
- use the sampled error probabilities $(\vec{e}'_0, \vec{e}'_1)$ to generate the linear transformation $\mathbf{C}'_k$
- apply the sampled correction matrix $\mathbf{C}'_l$ to m'_k to obtain the read-out mitigated estimator $\widetilde{m}'_k$

The resulting ensemble of measurements can be used directly to estimate expectation values and confidence intervals as described above. In this way, we avoid having to explicitly construct the variance of the correction matrix $\mathbf{C}'_l$ using maximum likelihood estimation and then propagating the error perturbatively to arbitrary observables as done in [69].

6.6.3 Zero-noise-extrapolation

For observables like inversion probability, we adopt the procedure developed in [69]. For entanglement observables we adopt a two point shifted exponential extrapolation that we briefly describe here. We denote the entanglement observable as $\langle F[O]\rangle^{(L)}(r)$ where L is the number of samples used and r denotes the noise level of the circuit, proportional to the number of CNOT gates in the circuit. We first note that in the case of very high noise levels, denoted here with $\langle F[O]\rangle(r \to \infty)$ the density matrix corresponds to the maximally mixed state given by $1/4$. Therefore, the concurrence in this case is $-1/2$ and the pair entanglement saturates to 2.

Using an estimate for the large noise expected value $\langle F[O]\rangle(r \to \infty)$, we can then consider a simple exponential extrapolation of the form

$$\langle F[O]\rangle^{(L)}(r) - \langle F[O]\rangle(r \to \infty) = A_F^{(L)} e^{-\alpha r}\,, \tag{6.71}$$

with α and $A_F^{(L)}$ the parameters of the model which can be obtain using results at two different noise levels r and r'. The zero-noise extrapolated result in this model corresponds to the limit $r \to 0$ and is given simply by the estimated $A_F^{(L)}$. More explicitly this becomes

$$A_F^{(L)} = \langle F[O]\rangle^{(L)}(r) \left(\frac{\langle F[O]\rangle^{(L)}(r')}{\langle F[O]\rangle^{(L)}(r)}\right)^{r/(r-r')}\,, \tag{6.72}$$

and the zero noise extrapolated observable is

$$\langle F[O]\rangle^{(L)}(0) = A_F^{(L)} + \langle F[O]\rangle(r \to \infty)\,. \tag{6.73}$$

Finally, the estimated statistical error is obtained by calculating the standard deviation of the L copies as above.

6.7 Additional data for concurrence and entanglement entropy

Here we show the full set of results for both entanglement entropy and concurrence for all the other pairs of qubits not shown earlier. We denote with a magenta triangle, data-points that fall below zero for the entropy as in Figure 6.9.

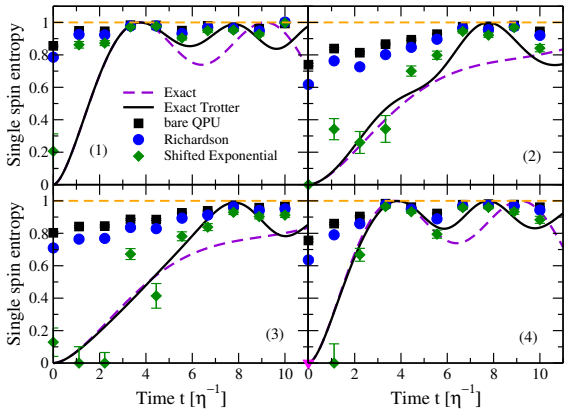


Figure 6.10: Single spin entanglement entropy for all four neutrinos. Black squares are bare results obtained from the QPU, the blue circles are obtained using Richardson extrapolation and the green diamonds correspond to the results obtained from a shifted exponential extrapolation using the maximum value of the entropy (dashed orange line). The magenta triangle indicates a mitigated result with shifted exponential extrapolation below zero within errorbars.

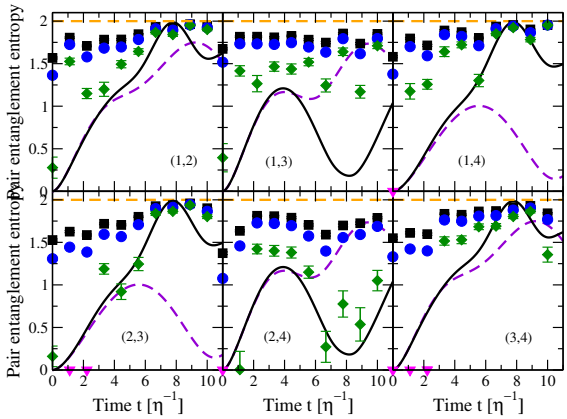


Figure 6.11: Pair entanglement entropy for all pairs of neutrinos. Black squares are bare results obtained from the QPU, red triangles are results obtained by amplifying the noise to $\epsilon/\epsilon_0 = 3$, the blue circles are obtained using Richardson extrapolation and the green diamonds correspond to the results obtained from a shifted exponential extrapolation using the maximum value of the entropy (indicated as a dashed orange line). The magenta triangle point are mitigated results with shifted exponential extrapolation below zero within errorbars.

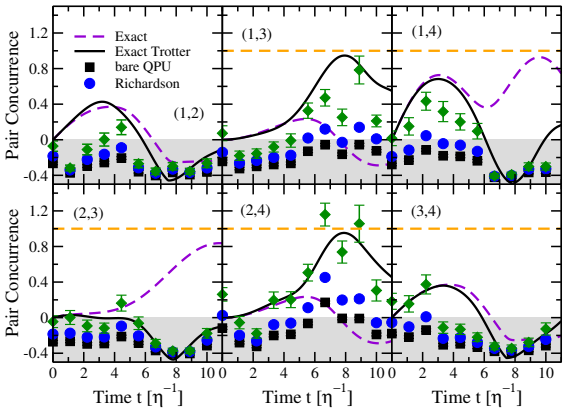


Figure 6.12: Entanglement concurrence for all the pairs of qubits. The maximum value for the concurrence is shown as a dashed orange line.

6.8 Conclusion

In this chapter, we presented the first digital quantum simulation of the flavor dynamics in collective neutrino oscillations using current quantum technology. The results reported for the evolution of flavor and entanglement properties of a system with $N = 4$ neutrino amplitudes show that current quantum devices based on superconducting qubits are starting to become a viable option for studying out-of-equilibrium dynamics of interacting many-body systems. The reduced fidelity in the results obtained here, compared to the simulations reported previously in [69] employing the same quantum processor and a comparable number of entangling gates, points to the importance of controlling unitary errors associated with the imperfect implementation of arbitrary single-qubit rotations (on average $< 1\%$ for the device used in both works). In future work we plan to explore the use of more advanced error mitigation strategies, such as Pauli twirling [84] or symmetry protection [81], to achieve a better overall fidelity.

We showed the zero-noise error extrapolation using a shifted Gaussian ansatz to be remarkably efficient in predicting the expected error-free estimator of observables. Given

the large circuits employed in this section, past experience with zero-noise extrapolations (see e.g. [70, 69]) suggest the exponential ansatz to be appropriate due to the large noise rates, and we find it to indeed outperforms Richardson extrapolation in this regime. Using the pair concurrence together with the entropy provides a robust way to detect entanglement even in the presence of substantial noise, like in the results shown here. We expect these insights, and the mapping of the neutrino evolution problem into a swap network, to prove very valuable in future explorations of out-of-equilibrium neutrino dynamics with near-term, noisy, quantum devices.

CHAPTER 7

QUANTUM CIRCUIT SQUEEZING ALGORITHM

7.1 Introduction

The quantum circuit squeezing algorithm (QCSA) is a novel quantum algorithm developed in this thesis which compresses the depth of a quantum circuit, with the trade-off of increased qubit size and required runs. The algorithm relies on two fundamental properties of maximally entangled states: the ricochet property and entanglement swapping. QCSA uses the combination of these two properties to replace horizontal lines of gates on a set of qubits to a vertical line of gates on a larger set of qubits.

7.2 Maximal Entanglement

QCSA relies extensively on maximally entangled states, a state whose entanglement entropy is maximal. A state $|\psi\rangle$ is maximally entangled if the partial trace of its density matrix $\rho = |\psi\rangle \langle\psi|$ is equal to a multiple of the identity I; mathematically

$$\mathrm{Tr}_A \rho_{AB} = cI_B, \tag{7.1}$$

where c is a constant. The partial trace can be defined as follows: Let $\{|a_i\rangle\}$ and $\{|b_i\rangle\}$ be the bases of Hilbert spaces H_A and H_B, respectively. Let $\rho_{AB} \in H_A \otimes H_B$ be a density matrix which can be decomposed as

$$\rho_{AB} = \sum_{ijkl} c_{ijkl} \, |a_i\rangle \, \langle a_j| \otimes |b_k\rangle \langle b_l| \,, \tag{7.2}$$

where c_{ijkl} are constants. Then the partial trace over A is given by

$$\mathrm{Tr}_A(\rho_{AB}) = \sum_{ijklm} c_{ijkl} \langle a_m|a_i\rangle \left\langle a_j \middle| a_m \right\rangle \otimes |b_k\rangle \langle b_l| \tag{7.3}$$

$$= \sum_{ijklm} c_{ijkl} \left\langle a_j \middle| a_m \right\rangle \langle a_m | a_i \rangle \otimes |b_k\rangle \langle b_l| \tag{7.4}$$

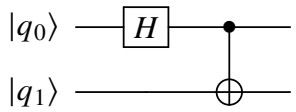
$$= \sum_{ijkl} c_{ijkl} \left\langle a_j \middle| a_i \right\rangle \otimes |b_k\rangle \langle b_l| \,. \tag{7.5}$$

One orthonormal set of maximally entangled states for two qubits, is the Bell state basis. It contains the following four states

$$\left|\phi^{\pm}\right\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle \pm |11\rangle\right) \tag{7.6}$$

$$\left|\psi^{\pm}\right\rangle = \frac{1}{\sqrt{2}} \left(|01\rangle \pm |10\rangle\right), \tag{7.7}$$

which can be prepared on a quantum computer with the following short-depth circuit



(7.8)

where the initial state of the pair of qubits $|q_0 q_1\rangle$ leads to the different Bell states as follows

$$|00\rangle \rightarrow \left|\phi^{+}\right\rangle, \tag{7.9}$$

$$|10\rangle \rightarrow |\phi^{-}\rangle, \tag{7.10}$$

$$|01\rangle \rightarrow \left|\psi^{+}\right\rangle, \tag{7.11}$$

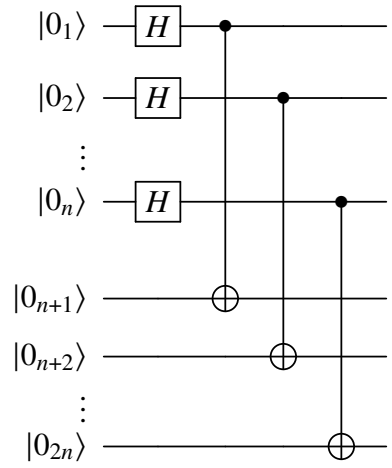
$$|11\rangle \rightarrow |\psi^{-}\rangle \,. \tag{7.12}$$

Each Bell state is maximally entangled because the partial trace of its density matrix is $I/2$. Note that preparing any of the initial states $|q_0 q_1\rangle$ listed above does not increase the depth of the circuit as their preparation requires at most the application of a column of two X gates, the top of which can be combined with the H gate to form a single-qubit gate and the bottom of which can be executed in parallel with this new single-qubit gate. The notion

of a maximally entangled state can be expanded to n qubits. One such state is what we'll call the generalized Bell state $\left|\phi_n^+\right\rangle$ which we'll define to be

$$\left|\phi_n^+\right\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in h_n} |xx\rangle \,, \tag{7.13}$$

where $h_n = \{x \mid l(x) = n\}$ is the set of all bitstrings of length n. The states formed from this set $H_n = \{|x\rangle \mid x \in h_n\}$, which we'll call the bit-string states, are an orthonormal basis for the Hilbert space of n-qubits. The state $\left|\phi_n^+\right\rangle$ can be prepared with the following circuit of depth two:



(7.14)

where the column of n Hadamard gates H places the top half of the qubits in the equal superposition of the states of H_n while the ladder of CNOT gates "copies" each state of the top half of the qubits to the state of the bottom half, maximally entangling the two sets of qubits. The circuit is depth two because all of the CNOT gates can be run in parallel as they each act on disjoint pairs of qubits. Note that this is simply n copies of the quantum circuit that prepares $|\phi^+\rangle$, (7.8) with $|q_0 q_1\rangle = |00\rangle$. The proof of the maximum entanglement of $\left|\phi_n^+\right\rangle$ is given below:

$$\mathrm{Tr}_A \left(\left|\phi^+\right\rangle \left\langle\phi^+\right|\right) = \mathrm{Tr}_A \left(\frac{1}{2^n} \sum_{x,y \in h_n} |xx\rangle \langle yy| \right)$$

$$
\begin{aligned}
&= \frac{1}{2^n} \sum_{x,y \in h_n} \langle x|y\rangle \, |x\rangle \langle y| \\
&= \frac{1}{2^n} \sum_{x,y \in h_n} \delta_{xy} \, |x\rangle \langle y| \\
&= \frac{1}{2^n} \sum_{x \in h_n} |x\rangle \langle x| \\
&= \frac{1}{2^n} I. \qquad (7.15)
\end{aligned}
$$

The maximally entangled states introduced here serve as the building blocks for QCSA. We will now discuss the two properties of maximally entangled states that will be used to build the algorithm.

7.3 Ricochet Property

The first property of maximally entangled states that will be used in QCSA is the Ricochet property. It states that, for any n-qubit gate A, the following equality holds

$$
(A \otimes I) \left|\phi_n^+\right\rangle = (I \otimes A^T) \left|\phi_n^+\right\rangle, \qquad (7.16)
$$

where one will recall $\left|\phi_n^+\right\rangle$ to be our previously defined (7.13) maximally entangled state for n qubits. The ricochet property can be proven as follows: First, write A in terms of the orthonormal basis of n-qubits H_n (which was defined as the set of all possible bit-string states of length n)

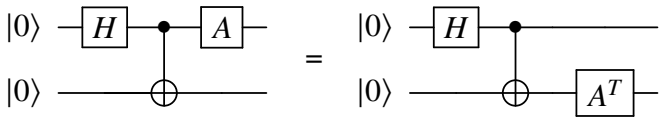
$$
A = \sum_{i,j \in h_n}^{n} a_{ij} \, |i\rangle \langle j| , \qquad (7.17)
$$

where a_{ij} are the matrix element of A. Then, plug this rewriting of A into the left hand side (LHS) of the ricochet property definition (7.16)

$$
(A \otimes I) \left|\phi_n^+\right\rangle = \frac{1}{\sqrt{n}} \sum_{ijk}^{n} a_{ij} (|i\rangle \langle j| \otimes I) \, |kk\rangle \qquad (7.18)
$$

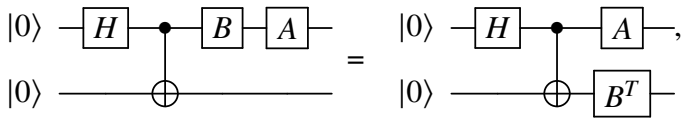
$$= \frac{1}{\sqrt{n}} \sum_{ij}^{n} a_{ij} |ij\rangle$$
$$= \frac{1}{\sqrt{n}} \sum_{ijk}^{n} a_{ij} (I \otimes |j\rangle \langle i|) |kk\rangle$$
$$= \left(I \otimes A^T\right) \left|\phi_n^+\right\rangle . \tag{7.19}$$

The ricochet property for single-qubit gates can be expressed via quantum circuits as:



(7.20)

The key insight here is that one can use this property to change two gates applied in series on a single qubit to two gates applied in parallel on two qubits. This process, which we'll call squeezing, can be seen in terms of quantum circuits below

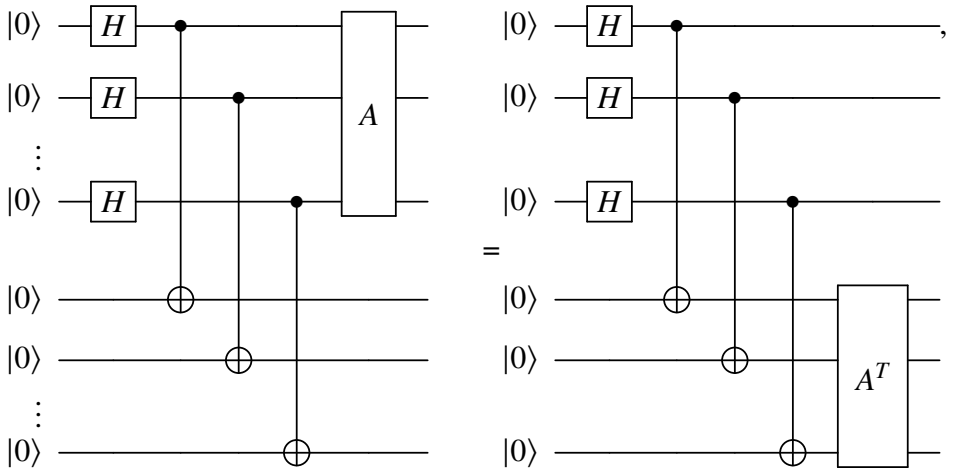


(7.21)

or, mathematically

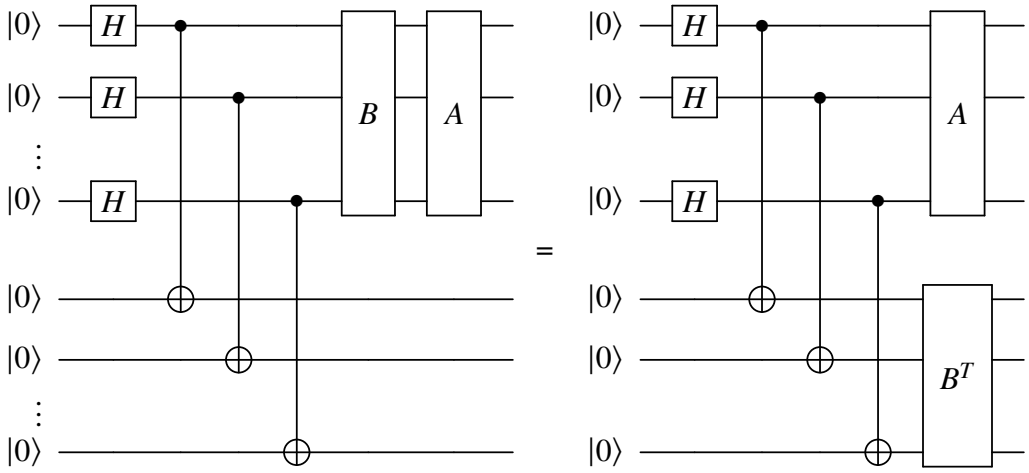
$$\begin{aligned} (AB \otimes I) \left|\phi_n^+\right\rangle &= (A \otimes I) \left[(B \otimes I) \left|\phi_n^+\right\rangle\right] \\ &= (A \otimes I) \left[(I \otimes B^T) \left|\phi_n^+\right\rangle\right] \\ &= (A \otimes B^T) \left|\phi_n^+\right\rangle . \end{aligned} \tag{7.22}$$

The ricochet property for the n-qubit case, using $\left|\phi_n^+\right\rangle$ can be visualized in terms of quantum circuits as follows



(7.23)

and the squeezing process can additionally be visualized as



(7.24)

So far, we've only been able to squeeze (reduce) our circuit depth by a factor 2. In order to do better, we'll need to introduce the second property of maximally entangled states upon which QCSA relies, entanglement swapping.

7.4 Entanglement swapping

The property of maximally entangled states that allows one to extend the benefits of the ricochet property to be able to squeeze the depth of a circuit by more than half is called entanglement swapping. To express it mathematically, let me first introduce the notation $\left|\phi^n_{ab}\right\rangle$ to mean that the generalized Bell state $\left|\phi^+_n\right\rangle$ (7.13) is applied to qubit sets a and b; that is

$$\left|\phi^n_{ab}\right\rangle = \frac{1}{\sqrt{2^n}} \sum_{x \in h_n} |x_a x_b\rangle \,, \tag{7.25}$$

where $h_n = \{x \mid l(x) = n\}$ is the set of all bitstrings of length n, while a and b are sets of n-qubits. For example, for $N = 2$, $a = \{0, 1\}$, and $b = \{2, 3\}$, we would have $h_n = \{00, 01, 10, 11\}$ and the generalized Bell state would be

$$\left|\phi^n_{ab}\right\rangle = \frac{1}{\sqrt{2^2}} \sum_{x \in h_2} |x_{\{0,1\}} x_{\{2,3\}}\rangle = \frac{1}{2} \left(|0000\rangle + |0101\rangle + |1010\rangle + |1111\rangle\right) . \tag{7.26}$$

With this notation in hand, the entanglement swapping property can then be stated as follows:

$$\left\langle \phi^n_{bc} \middle| \phi^n_{ab} \phi^n_{cd} \right\rangle = \frac{1}{2^n} \left|\phi^n_{ad}\right\rangle , \tag{7.27}$$

which can be proven by inserting into it the definition of the generalized Bell state (7.25), which yields

$$\begin{aligned} \left\langle \phi^n_{bc} \middle| \phi^n_{ab} \phi^n_{cd} \right\rangle &= \left(\frac{1}{\sqrt{2^n}} \sum_{x \in h_n} \langle x_b x_c| \right) \left(\frac{1}{2^n} \sum_{y,z \in h_n} |y_a y_b z_c z_d\rangle \right) \\ &= \frac{1}{\sqrt{2^{3n/2}}} \sum_{x \in h_n} |y_a\rangle \langle x_b|y_b\rangle \langle x_c|z_c\rangle |z_d\rangle \\ &= \frac{1}{\sqrt{2^{3n/2}}} \sum_{x \in h_n} \delta_{xy} \delta_{xz} |y_a\rangle |z_d\rangle \end{aligned} \tag{7.28}$$

$$= \frac{1}{2^n}\left(\frac{1}{\sqrt{2^n}}\sum_{x\in h_n}|x_a x_d\rangle\right)$$
$$= \frac{1}{2^n}\left|\phi^n_{ad}\right\rangle. \tag{7.29}$$

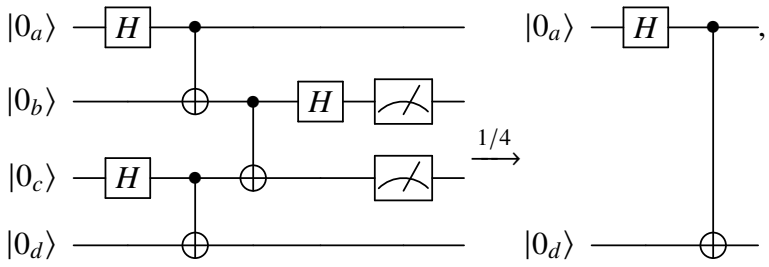
Taking the complex conjugate squared of both sides of the entanglement swapping property (7.27)

$$\left|\left\langle\phi^n_{bc}\middle|\phi^n_{ab}\phi^n_{cd}\right\rangle\right|^2 = \frac{1}{4^n}\left\langle\phi^n_{ad}\middle|\phi^n_{ad}\right\rangle = \frac{1}{4^n}, \tag{7.30}$$

implies that the probability of measuring qubit sets b and c to be in the state $\left|\phi^+_n\right\rangle$ is $1/4^n$. One can test whether qubit sets b and c are in the state $\left|\phi^+_n\right\rangle$ by applying $(U^n_{bc})^\dagger$ to said qubits, measuring them, and checking if they were both measured to be in the all zero state $|0\rangle^{\otimes n}$. Here, we've defined U^n_{ab} as the quantum gate that takes qubit sets a and b from all zero states to the generalized Bell state $\left|\phi^n_{ab}\right\rangle$; that is

$$U_n\,|0\rangle^{\otimes n}_a\,|0\rangle^{\otimes n}_b = \left|\phi^n_{ab}\right\rangle. \tag{7.31}$$

To get a better feel for the entanglement swapping property, we show here the property in terms of quantum circuits for the $n = 1$ case:



, (7.32)

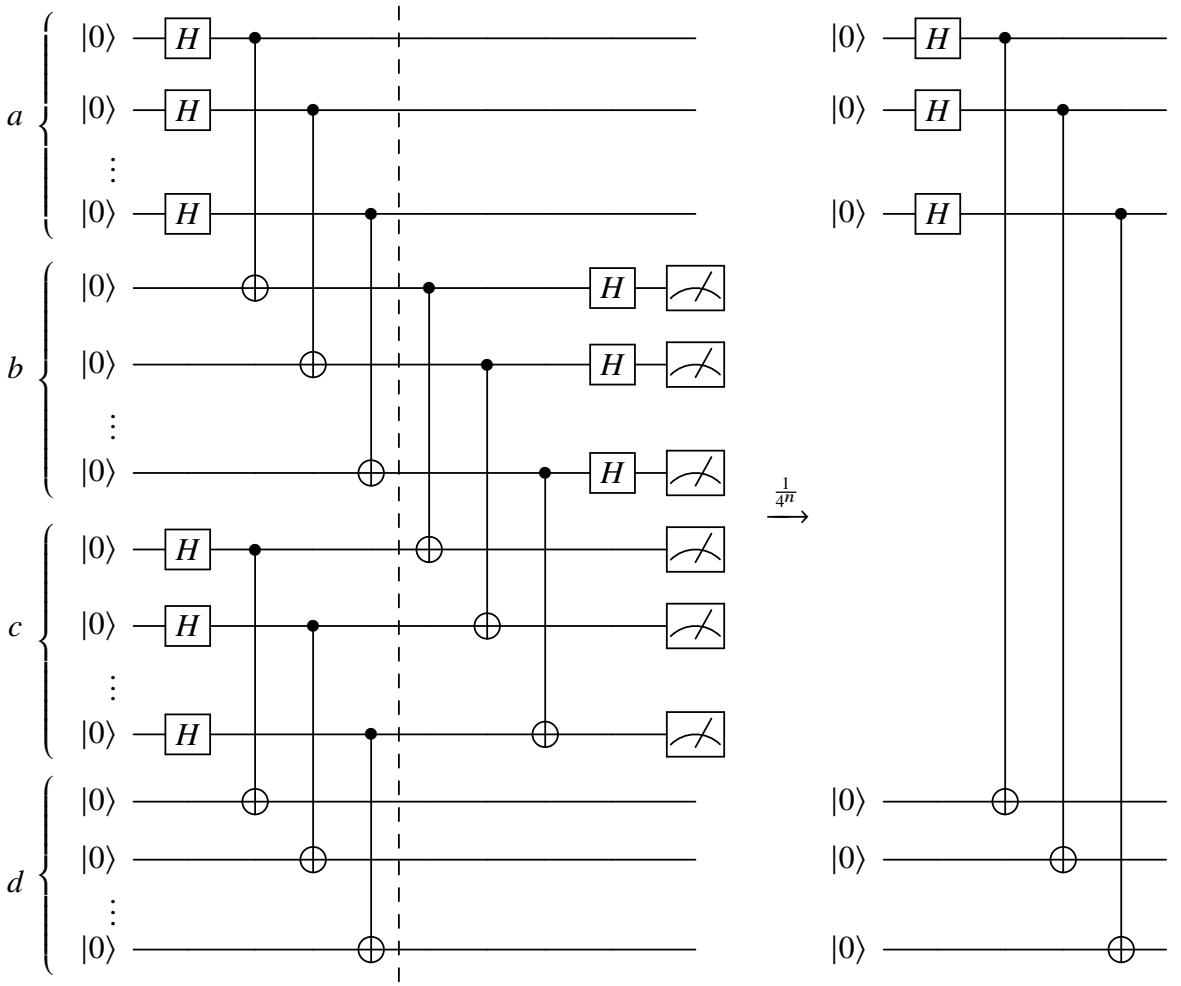
whose mathematical description matches that of the $n = 1$ case of the definition of the property (7.27)

$$\langle 00|\,(U^n_{bc})^\dagger U^n_{ab} U^n_{bc}\,|0000\rangle = \frac{1}{2^n}\left|\phi^n_{ad}\right\rangle \tag{7.33}$$

$$\left\langle \phi^n_{bc} \middle| \phi^n_{ab} \phi^n_{cd} \right\rangle = \frac{1}{2^n} \left| \phi^n_{ad} \right\rangle , \tag{7.34}$$

where we've used the definition 7.31. Here the arrow with $1/4$ above it from 7.32 implies that qubits b and c were measured to be 00 with probability $1/4$. Additionally, qubits b and c have been discarded in the circuit to the right of the arrow as they have been measured and are therefore no longer relevant. Note that, in the $n = 1$ case considered above U^n_{xy} is simply the quantum circuit that prepares $|\phi^+\rangle$ given in 7.8 (with $x = q_0$ and $y = q_1$); that is, a Hadamard applied to the top qubit x and a CNOT with the qubits x and y being the control and target qubits, respectively. We'll now consider the arbitrary n case (7.27). Here, U^n_{xy} is the quantum circuit given in (7.14). This case is given in the quantum circuit representation below, in which the arrow with a $\frac{1}{4^n}$ above it implies that the qubit sets b

and c were measured to be in the all zero state $|0\ldots0\rangle$:



(7.35)

7.5 Dicke Subspace Modification

The quantum circuit squeezing algorithm can be modified for circuits that preserve Hamming weight (for example, ansatzes in second quantization for systems that preserve particle number) so that the increase in additional measurements is reduced. To describe this modification, let us first define what we'll call the Dicke-Bell state $|\phi_k^n\rangle$, which can be

thought of as a maximally entangled state in the sub-space of the Hilbert space for n qubits containing only the bit-string states that have a Hamming weight of k. Mathematically

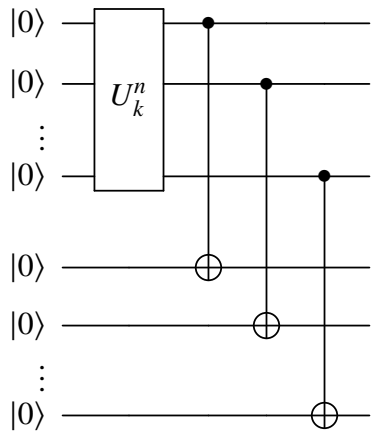
$$\left|\phi_k^n\right\rangle = \frac{1}{\sqrt{\binom{n}{k}}} \sum_{x \in h_k^n} |xx\rangle\,, \tag{7.36}$$

where $h_k^n = \{x \mid \mathrm{l}(x) = n, \mathrm{wt}(x) = k\}$; that is the set of all bit-strings x with length n and Hamming weight k. For example,

$$\left|\phi_2^4\right\rangle = \frac{1}{\sqrt{6}}(|1100\rangle \otimes |1100\rangle + |1010\rangle \otimes |1010\rangle + |1001\rangle \otimes |1001\rangle \tag{7.37}$$

$$+ \ |0110\rangle \otimes |0110\rangle + |0101\rangle \otimes |0101\rangle + |0011\rangle \otimes |0011\rangle). \tag{7.38}$$

It is analogous to the state $\left|\phi_n^+\right\rangle$ defined in (7.13) with x being drawn from the set h_k^n instead of h_n. The states created from the latter, $H_n = \{|x\rangle \ \ |x \in h_n\}$, are an orthonormal basis for the Hilbert space of n-qubits. This is as opposed to the states created from the former, $H_k^n = \{|x\rangle \ \ |x \in h_k^n\}$, which is an orthonormal basis for H_k^n, the subset of the Hilbert space for n-qubits containing only quantum states formed from bit-string states with a Hamming weight of k. The Dicke-Bell state $|\phi_k^n\rangle$ can be formed from the following circuit



(7.39)

where U_k^n takes the n-qubit all-zero state to the n, k Dicke state; that is

$$U_k^n \, |0\rangle^{\otimes n} = \, |D_k^n\rangle\,, \tag{7.40}$$

where the Dicke state $|D_k^n\rangle$ is defined as the equal superposition of all bit-string states of length n and Hamming weight k; that is

$$|D_k^n\rangle = \frac{1}{\sqrt{\binom{n}{k}}} \sum_{x \in h_k^n} |xx\rangle \,. \tag{7.41}$$

The reason we've named the resulting state the Dicke-Bell state is now revealed; it is the entanglement of a Dicke-state into a maximally entangled state for a constant Hamming weight subspace. Finding a short-depth circuit decomposition for U_k^n (to prepare a Dicke state) is an active area of research. It has been shown how to construct U_k^n with $O(n)$ depth and $O(kn)$ gates ([9] and [2]). While linear depth is certainly not as good as the constant two depth circuits (7.14) we shall see that using the subspace modification (which requires the preparation of Dicke states) provides a significant advantage in terms of amount of shots (runs of quantum circuits) required. Additionally, more efficient preparation methods may be discovered, including our novel method considered in chapter 8.

It can be shown that the ricochet property still holds for $|\phi_k^n\rangle$. That is, for any n by n matrix A_k that preserves the Hamming weight (k) of any state upon which it acts, the following equality holds

$$(A \otimes I) \left|\phi_k^n\right\rangle = (I \otimes A^T) \left|\phi_k^n\right\rangle . \tag{7.42}$$

It can be proven analogously to the proof of the original ricochet property (7.18) by writing A in terms of the orthonormal basis h_k^n

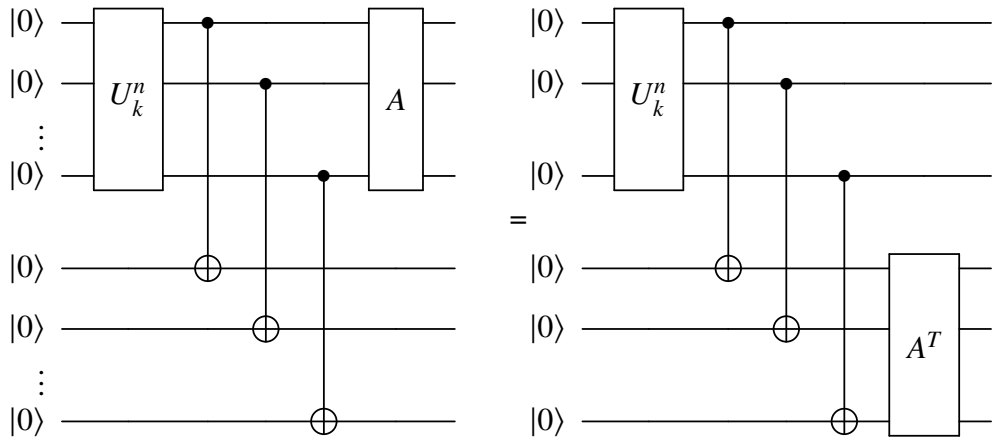
$$A_k = \sum_{x,y \in h_k^n} a_{xy} \, |x\rangle \langle y| \,, \tag{7.43}$$

and calculating

$$(A_k \otimes I) \left|\phi_k^n\right\rangle = \frac{1}{\sqrt{\binom{n}{k}}} \sum_{xyz} a_{xy} (|x\rangle \langle y| \otimes I) \, |zz\rangle \tag{7.44}$$

$$
\begin{aligned}
&= \frac{1}{\sqrt{\binom{n}{k}}} \sum_{xy} a_{xy} \left|xy\right\rangle \\
&= \frac{1}{\sqrt{\binom{n}{k}}} \sum_{xyz} a_{xy} (I \otimes \left|y\right\rangle \left\langle x\right|) \left|zz\right\rangle \\
&= \left(I \otimes A_k^T\right) \left|\phi_k^n\right\rangle ,
\end{aligned}
$$

where x, y, z sum over the set h_k^n. The quantum circuit representation of this ricochet property is given below



(7.45)

Because A_k preserves Hamming weight, it sees the subspace H_k^n the same way that A sees H^n, as the entire space that it can explore. In other words, H_k^n is closed under the application of A_k just as H^n is closed under the application of A. Mathematically, $\forall x \in H_k^n$, $A_k x \in H_k^n$ just as $\forall x \in H_n$, $Ax \in H_n$.

Entanglement swapping also holds for $\left|\phi_k^n\right\rangle$. To explain, let me first introduce the notation $\left|\phi_{ij}^{nk}\right\rangle$ to mean that the Dicke-Bell state $\left|\phi_k^n\right\rangle$ is applied to qubits i and j; that is

$$
\left|\phi_{ij}^{nk}\right\rangle = \frac{1}{\sqrt{\binom{n}{k}}} \sum_{x \in h_k^n} \left|x_i x_j\right\rangle . \tag{7.46}
$$

Then, the analogous entanglement swapping property is

$$\left\langle \phi_{bc}^{nk} \middle| \phi_{ab}^{nk} \phi_{cd}^{nk} \right\rangle = \frac{1}{\binom{n}{k}} \left| \phi_{ad}^{nk} \right\rangle , , \tag{7.47}$$

which can be proven by inserting into it the definition of the Dicke-Bell state (7.46) which yields

$$\left\langle \phi_{bc}^{nk} \middle| \phi_{ab}^{nk} \phi_{cd}^{nk} \right\rangle = \left(\frac{1}{\sqrt{\binom{n}{k}}} \sum_{x \in h_k^n} \langle x_b x_c | \right) \left(\frac{1}{\binom{n}{k}} \sum_{y,z \in h_k^n} |y_a y_b z_c z_d\rangle \right) \tag{7.48}$$

$$= \frac{1}{\binom{n}{k}^{3/2}} \sum_{x,y,z \in h_k^n} |y_a\rangle \langle x_b | y_b\rangle \langle x_c | y_c\rangle |z_d\rangle \tag{7.49}$$

$$= \frac{1}{\binom{n}{k}^{3/2}} \sum_{x,y,z \in h_k^n} \delta_{xy} \delta_{xz} |y_a\rangle |z_d\rangle \tag{7.50}$$

$$= \frac{1}{\binom{n}{k}} \left(\frac{1}{\sqrt{\binom{n}{k}}} \sum_{x \in h_k^n} |x_a\rangle |x_d\rangle \right) \tag{7.51}$$

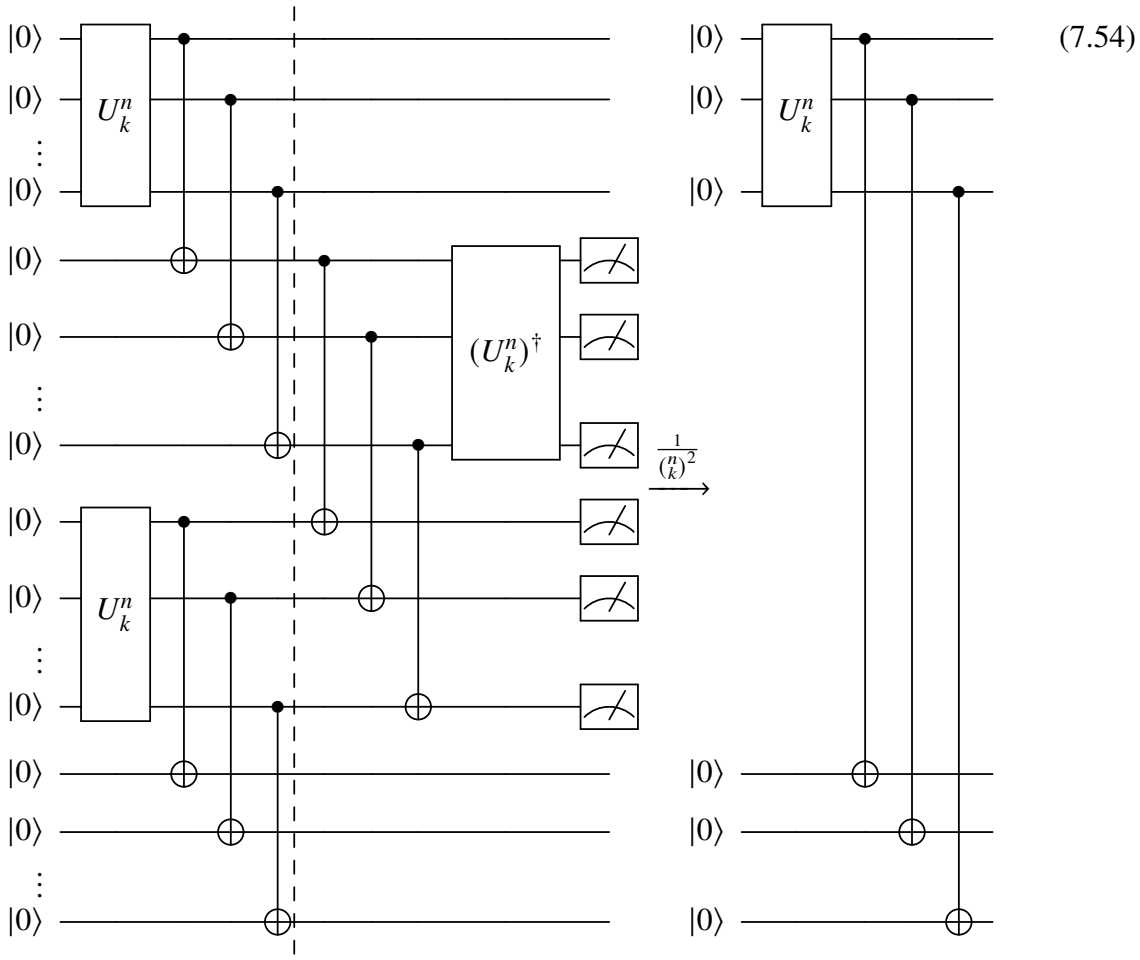
$$= \frac{1}{\binom{n}{k}} \left| \phi_{ad}^{nk} \right\rangle . \tag{7.52}$$

Taking the complex conjugate squared of both sides of the entanglement swapping property (7.27)

$$\left| \left\langle \phi_{bc}^{nk} \middle| \phi_{ab}^{nk} \phi_{cd}^{nk} \right\rangle \right|^2 = \frac{1}{\binom{n}{k}^2} \left\langle \phi_{ad}^{nk} \middle| \phi_{ad}^{nk} \right\rangle = \frac{1}{\binom{n}{k}^2}, \tag{7.53}$$

implies that the probability of measuring qubit sets b and c to be in the state $\left| \phi_k^n \right\rangle$ is $\frac{1}{\binom{n}{k}^2}$. Note that the probability of success (every qubit in the qubit sets b and c was measured to be zero) has gone from scaling exponentially with n (7.27) to scaling linearly with n (7.47), the benefits of which will become clear soon. The entanglement swapping property can be viewed in terms of quantum circuits below, in which the arrow with a $1/\binom{n}{k}^2$ above it

means that the probability of equivalency is $1/\binom{n}{k}^2$. The dashed line is to delineate between the circuit between the section that entangles sets of qubits and the section that rotates the basis as to allow one to measure in the Dicke-Bell basis.



(7.54)

7.6 Entanglement Swapping Recursion

As we've seen, the procedure of entanglement swapping can be used to reduce the length of a circuit by the order of a factor of two. However, we can further reduce the circuit length by extending entanglement swapping in a recursive manner: First, we define

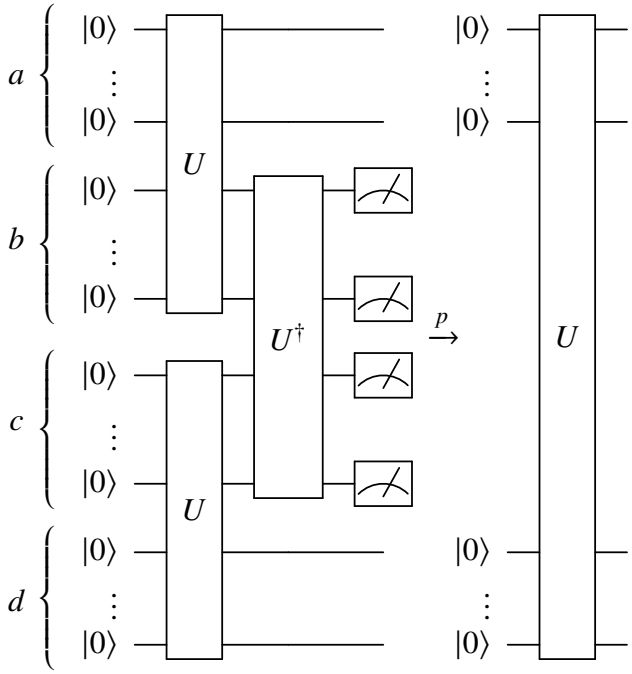
the base case:

$$\langle 0| U_{bc}^{\dagger} U_{ab} U_{cd} |0\rangle = \sqrt{p} U_{ad} |0\rangle , \tag{7.55}$$

where

$$U_{ab} |0\rangle = |\phi_{ab}\rangle , \tag{7.56}$$

with $|\phi_{ab}\rangle$ defined as a general maximally entangled state which is maximally entangled in some subspace H. We define it generally here so as to cover the previously explored cases: when $H = H^n$, the full Hilbert space of n qubits, and when $H = H_k^n$, the subspace of H^n restricted to states of Hamming weight k. (However, H can be any subspace of H^n.) The state formed by applying U_{ab} to the vacuum and the probability p are determined by H. For example, when $H = H^n$, we have $U_{ab} |0\rangle = |\phi_{ab}^n\rangle$ and $p = 1/4^n$. Meanwhile, when $H = H_k^n$, we have $U_{ab} |0\rangle = |\phi_{ab}^{nk}\rangle$ and $p = 1/\binom{n}{k}^2$. This base case can be represented by the following quantum circuit diagram.



(7.57)

The recursive extension of entanglement swapping can be stated mathematically as

$$\langle 0| \prod_{k=1}^{N-1} U^{\dagger}_{2k-1,2k} \prod_{k=1}^{N} U_{2k-2,2k-1} |0\rangle = \sqrt{p^{N-1}} U_{0,2N-1} |0\rangle \,, \tag{7.58}$$

where the indices of U refer to qubit sets. To prove this, we assume as the induction hypothesis, that (7.58) is true and proceed to show that the statement holds when we take $N \longrightarrow N+1$:

$$\langle 0| \prod_{k=1}^{N} U^{\dagger}_{2k-1,2k} \prod_{k=1}^{N+1} U_{2k-2,2k-1} |0\rangle \,; \tag{7.59}$$

pulling out the $k = N$ term from the first product and the $k = N$ and $k = N+1$ terms from the second product yields

$$\langle 0| \prod_{k=1}^{N-1} U^{\dagger}_{2k-1,2k} \prod_{k=1}^{N-1} U_{2k-2,2k-1} |0\rangle \langle 0| U^{\dagger}_{2N-1,2N} U_{2N-2,2N-1} U_{2N,2N+1} |0\rangle \,; \tag{7.60}$$

using the base case yields

$$\sqrt{p} \langle 0| \prod_{k=1}^{N-1} U^{\dagger}_{2k-1,2k} \prod_{k=1}^{N-1} U_{2k-2,2k-1} |0\rangle U_{2N-2,2N+1} |0\rangle \,; \tag{7.61}$$

relabeling qubit set $2N+1 \longrightarrow 2N-1$ (which is allowed as qubit set $2N-1$ has been measured) yields

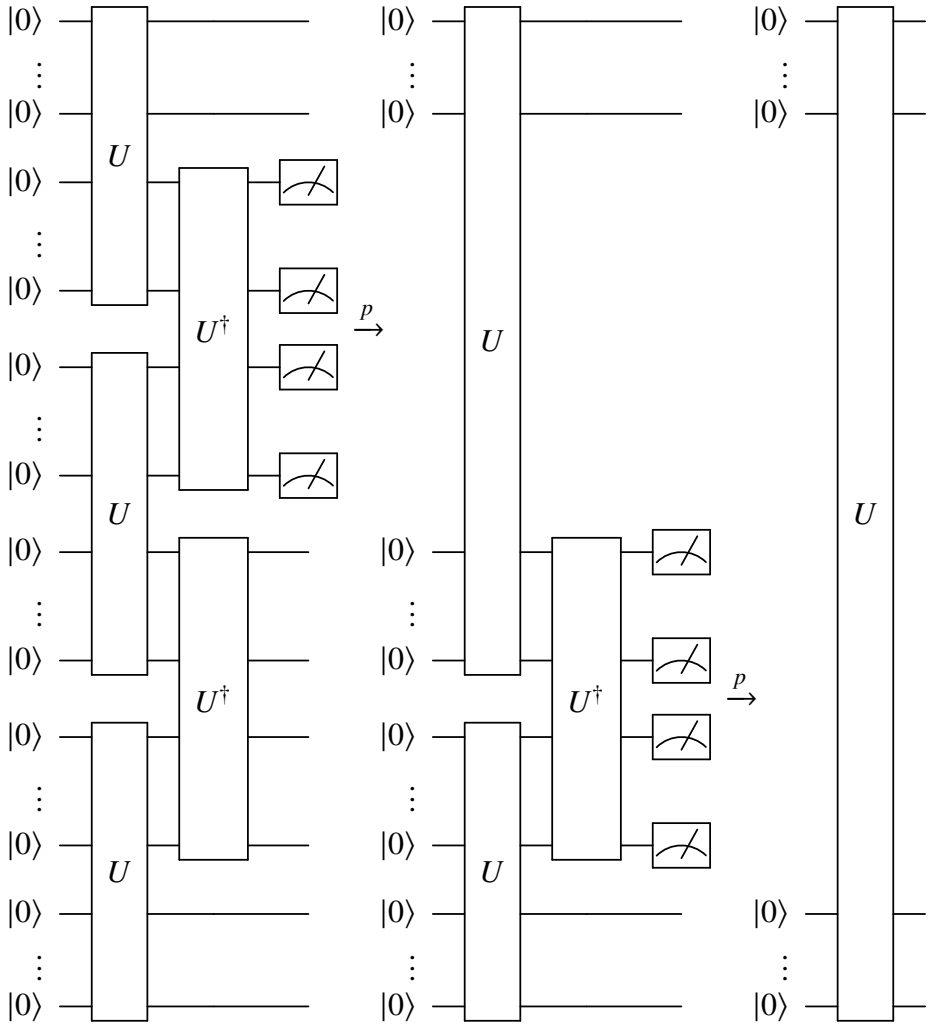
$$\langle 0| \prod_{k=1}^{N-1} U^{\dagger}_{2k-1,2k} \prod_{k=1}^{N} U_{2k-2,2k-1} |0\rangle \,; \tag{7.62}$$

applying the induction hypothesis gives

$$\sqrt{p^{N}} U_{0,2N-1} |0\rangle \tag{7.63}$$

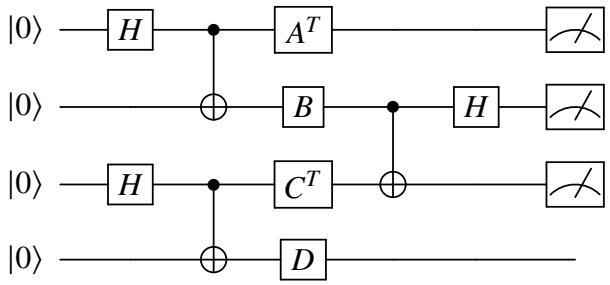
which completes the proof. The quantum circuit representation of entanglement swapping recursion for the case $N = 3$ is given below. The arrow with the p above it implies that

the following circuit is equivalent to the previous with probability p. Between the first and second circuit, the entanglement swapping property is applied amongst the first four sets of qubits while between the second and third circuit, the entanglement swapping property is applied amongst the remaining sets of qubits.



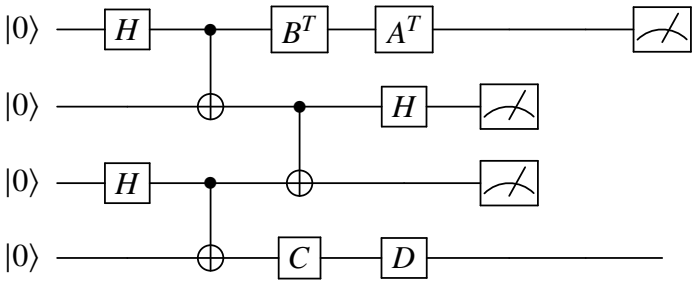
7.7 The Algorithm

We will now combine everything introduced thus far (ricochet property, entanglement swapping property, and entanglement swapping recursion) in order to construct the quantum circuit squeezing algorithm (QCSA). Before we give a formal definition, we'll walk through the algorithm for single-qubit gates and a single recursion step in quantum circuit form: Let A, B, C, D be arbitrary single-qubit gates. Then the QCSA circuit is given by



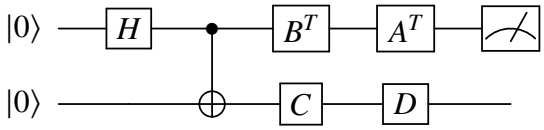
(7.65)

Since the qubit pairs $(0, 1)$ and $(2, 3)$ are each in the $|\phi^+\rangle$ state, one can apply the ricochet property to move the middle two gates (B and C^T) outward



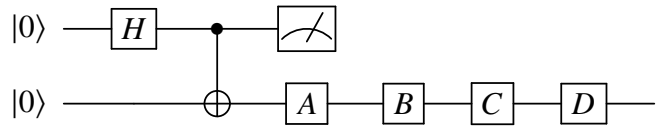
(7.66)

The entanglement swapping identity then tells us that, with a probability of $1/4$, measuring the middle two qubits will collapse them to the state $|00\rangle$, which would result in the following circuit



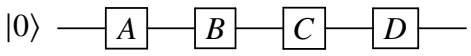
(7.67)

where the middle two qubits have been discarded. Since the two qubits left are in the state $|\phi^+\rangle$, the ricochet property can be applied again to move the top two gates (B^T and A^T) down to the bottom qubit.



(7.68)

One then measures the first qubit. With probability 1/2, one will measure 0 which implies that the second qubit is in the state $|0\rangle$, since the two qubits are in the entangled state $|\phi^+\rangle$. Thus, with probability 1/2, the above circuit is equivalent to the following circuit



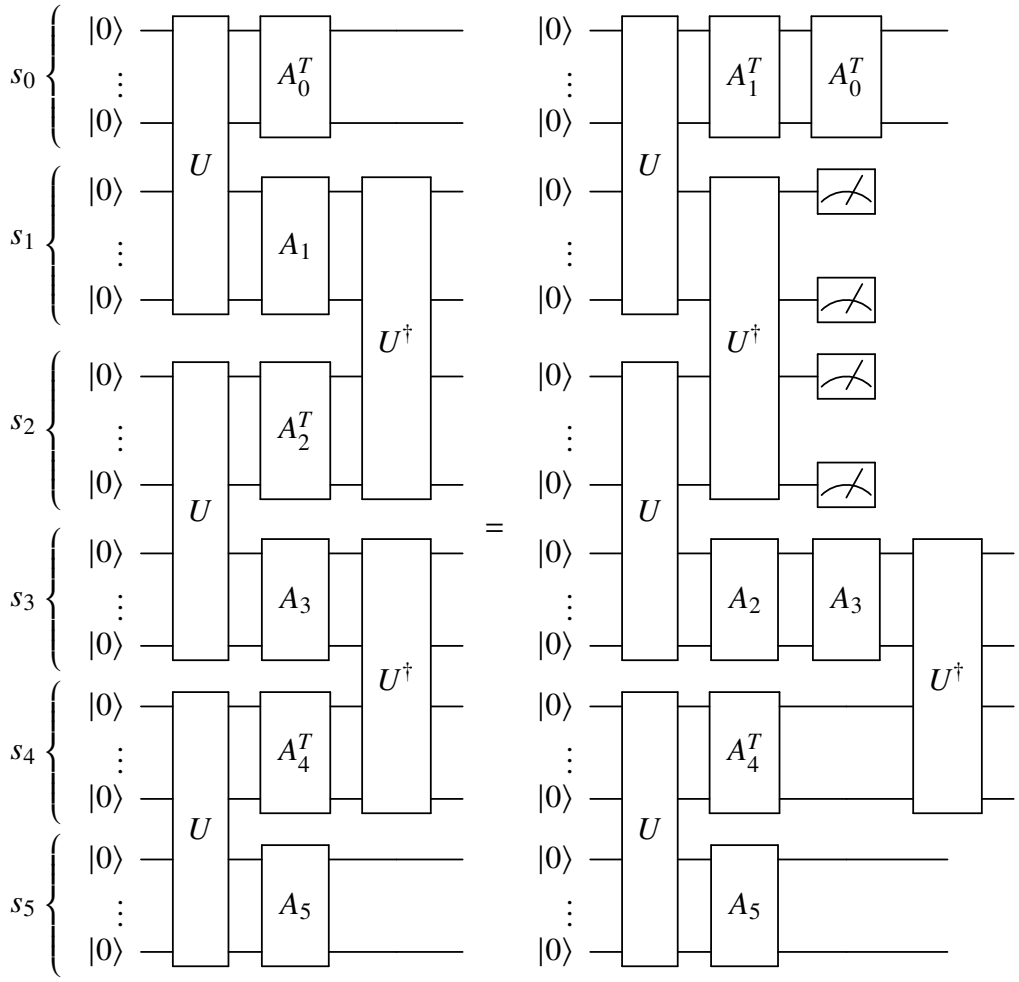
(7.69)

QCSA's benefit can now be understood: If one desires to run the circuit (7.69), one can instead run the squeezed circuit (7.65) which allows one to apply the four single-qubit gates A, B, C, D in parallel instead of series (with the trade-off of the usage of more qubits and the circuits only being equivalent one fourth of the time). While these two circuits have the same depth, we will see that for a large initial depth, the depth of the squeezed circuit can be decreased substantially through the use of QCSA. Since the final circuit (7.69) is only equivalent to the first circuit (7.65) with probability 1/8, one must run the first circuit eight times the number of runs one desires to run the first circuit. Assuming that these runs cannot be done in parallel, this increases the total run time by a factor of eight. However, we shall see later that the factor by which the run time increases need not necessarily scale exponentially with the depth of the original circuit.

We now walk through a more robust example of QCSA. Here we will use U and $|\Phi\rangle$ so as to keep the description of the algorithm general. That is, $|\Phi\rangle$ can refer to either the

generalized Bell state $\left|\phi_n^+\right\rangle$ (7.25) or the Dicke-Bell state $|\phi_k^n\rangle$ (7.36), with U being the operator that transforms the all zero state into the chosen version of $|\Phi\rangle$. The algorithm works the same for both cases, the only difference being the probability p of one each subsequent circuit being equivalent after measurement. However, this probability p will be given for both cases at each such step as we continue. In (7.70) we start on the left hand side (LHS) with an initial circuit which consists of six sets of qubits (s_0 through s_5). These six sets are entangled pairwise into three maximally entangled states $|\Phi_{01}\Phi_{23}\Phi_{45}\rangle$ via the application of $U_{01}U_{23}U_{45}$. Then we apply a column of the six gates that we actually wish to run (A_0 through A_5), transposing the even-indexed ones. Finally, we apply $U^\dagger$ to the inner two pairs of qubit sets (s_1,s_2) and (s_3,s_4). To get to the right hand side (RHS) of (7.70), we apply the ricochet property between pairs of qubit sets (s_0,s_1) and (s_2,s_3). Doing so ricochets A_1 from qubit set s_1 up to qubit set s_0 and ricochets A_2^T from qubit set s_2 down to qubit set s_3. This leaves space for $U_{12}^\dagger$ to be next to U_{01} and U_{23}, allowing us to apply

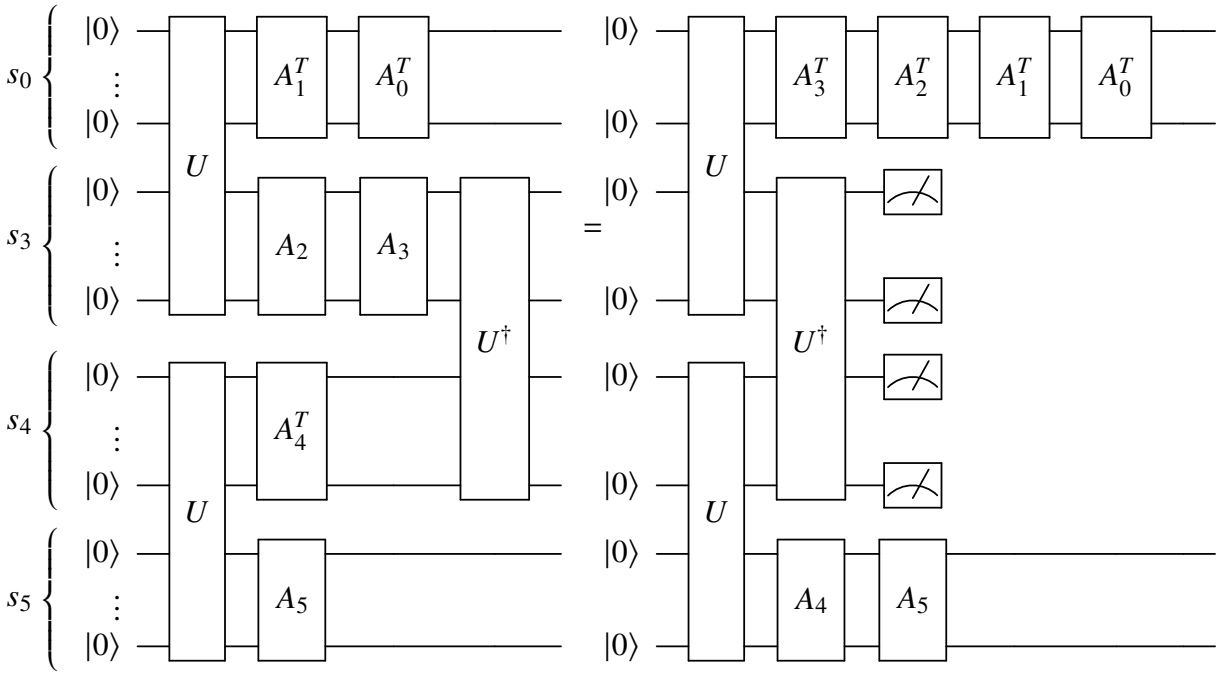
the entanglement swapping identity to the set of qubit sets $\{s_0, s_1, s_2, s_3\}$.



(7.70)

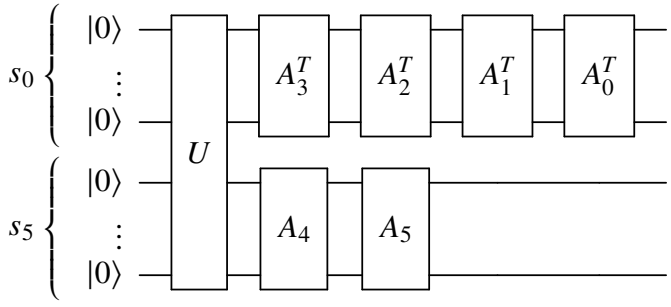
This means that the RHS of the circuit above (7.70) is equal to the LHS of the circuit below (7.71) with probability p. The probability p is given by $p = 1/4^n$ if U prepares $\left|\phi_n^+\right\rangle$ or $p = 1/\binom{n}{k}^2$ if U prepares $\left|\phi_k^n\right\rangle$. Here n is the number of qubits in each qubit set s_i and k is the Hamming weight that one can set as desired. To get to the LHS of (7.71) we apply the ricochet property between pairs of qubit set (s_0, s_3) and (s_4, s_5). Doing so ricochets A_2 and A_3 from qubit set s_3 up to qubit set s_0 and ricochets A_4^T from qubit set s_4 down to qubit set s_5. This leaves space for $U_{34}^\dagger$ to be next to U_{03} and U_{45}, allowing us to apply the

entanglement swapping identity on qubit sets $\{s_0, s_3, s_4, s_5\}$.



(7.71)

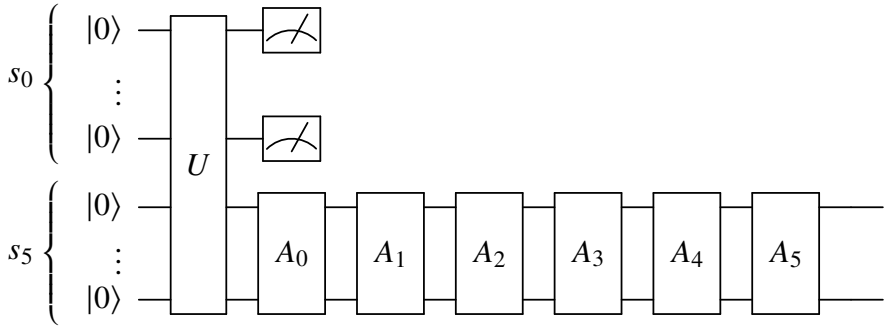
This means that the RHS of the circuit above (7.71) is equal to the LHS of the circuit below (7.72) with probability p, defined the same as before. To get to the LHS of (7.72) we apply the ricochet property between qubit set s_0 and s_5.



(7.72)

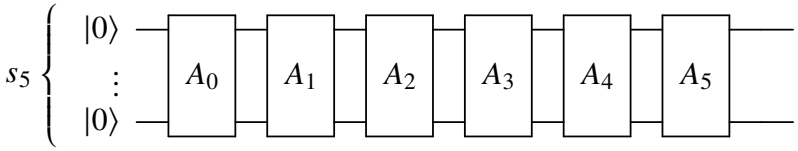
Applying the ricochet property between qubit sets q_0 and q_5 gives us the circuit below the

next circuit (7.73)



(7.73)

Finally, we measure the qubit set s_0. The circuit above (7.73) is equal to the circuit below (7.74) with probability q. If U prepares $\left|\phi_n^+\right\rangle$, then $q = 1/2^n$. If U prepares $|\phi_k^n\rangle$ then $q = 1/\binom{n}{k}$.



(7.74)

Working backwards, this implies that if one desires to run the circuit above (7.74), one can instead run the shorter depth squeezed circuit (7.74). All together, the probability p of these two circuits being equal is $p = 1/2^{5n}$ if U prepares $\left|\phi_n^+\right\rangle$ or $p = 1/\binom{n}{k}^5$ if U prepares $|\phi_k^n\rangle$.

We now describe the QCSA algorithm in full. The algorithm takes as an input, a depth n quantum circuit consisting of m qubits and an even integer d between 1 and $m + 1$. The algorithm returns as an output, a depth $\lceil n/d \rceil$ (squeezed) quantum circuit consisting of dm. To begin, one groups the n gate columns of the input circuit into $n \mod d$ groups of depth $\lceil n/d \rceil$ and $d - (n \mod d)$ groups of depth $\lfloor n/d \rfloor$. We label these d groups of gate columns $A_0, \ldots, A_{d-1}$. With this grouping, the algorithm is described via pseudo-code in Algorithm 7.1, below.

Algorithm 7.1 Quantum Circuit Squeezing Algorithm

Input: A quantum circuit consisting of m qubits (q_i for $i \in 0, \ldots, m-1$) and d groups of gate columns (A_i for $i \in 0, \ldots, n-1$). An even factor d by which the depth of the layers of gates will be shortened.

Output: Squeezed quantum circuit consisting of dm qubits and $\lceil n/d \rceil$ layers of gates.

success = False
while success = False **do**
  Initialize squeezed quantum circuit with dm qubits.
  for $0 \leq i < d/2$ **do**
    Apply U to qubit set $\{q_{2mi}, \ldots, q_{2m(i+1)-1}$
  end for
  for $0 \leq i < d$ **do**
    if i is even **then**
      apply A_i^T to the qubit set $\{q_{mi}, \ldots, q_{m(i+1)-1}\}$
    end if
    if i is odd **then**
      apply A_i to the qubit set $\{q_{(m(2i+1)}, \ldots, q_{m(2i+3)-1}\}$
    end if
  end for
  for $0 \leq i < d/2 - 1$ **do**
    apply $U^\dagger$ the to qubit set $\{q_{(m(2i+1)}, \ldots, q_{m(2i+3)-1}\}$
  end for
  for $0 \leq i < (d-1)m$ **do**
    Measure qubit q_i to bit c_i
  end for
  if $c_i = 0$ for all $i \in \{0, \ldots m(d-1) - 1$ **then**
    success = True
    Discard measured qubits and return circuit
  end if
end while

We will now prove QCSA (Algorithm 7.1) by induction. For this proof, we define the following parameters: let s_i (for any i) denote an m-qubit set; let A_i (for any i) denote a group of gate columns, and let d be an even positive integer. With this, QCSA for arbitrary

d can be stated mathematically as

$$\left(\prod_{k=1}^{d-2} \langle 0_{s_k}|\right)\left(\prod_{k=0}^{d/2-2} U^{\dagger}_{s_{2k+1},s_{2k+2}}\right)\left(\prod_{k=0}^{d/2-1} (A^T_{2k})_{s_{2k}}\right)$$
$$\times \left(\prod_{k=0}^{d/2-1} (A_{2k+1})_{s_{2k+1}}\right)\left(\prod_{k=0}^{d/2-1} U_{s_{2k},s_{2k+1}}\right)\left(\prod_{k=0}^{d-1} |0_{s_k}\rangle\right)$$
$$= \sqrt{p^{d/2-1}}\left(\prod_{k=0}^{d-1}(A_k)_{s_{2d-1}}\right) U_{s_0,s_{2d-1}} \left|0_{s_0} 0_{s_{d-1}}\right\rangle . \tag{7.75}$$

The base case ($d = 2$) for QCSA is

$$\left\langle 0_{s_1} 0_{s_2}\right| U^{\dagger}_{s_1 s_2} (A^T_0)_{s_0} (A_1)_{s_1} (A^T_2)_{s_2} (A_3)_{s_3} U_{s_0 s_1} U_{s_2 s_3} \left|0_{s_0} 0_{s_1} 0_{s_2} 0_{s_3}\right\rangle \tag{7.76}$$
$$= p(A^T_0 A^T_1)_{s_0} (A^T_3 A^T_2)_{s_3} U_{s_0 s_3} \left|0_{s_0} 0_{s_3}\right\rangle , \tag{7.77}$$

which can be proved as follows: first, we apply the ricochet property on qubit set pairs $\{s_0, s_1\}$ and $\{s_2, s_3\}$, which results in

$$U^{\dagger}_{s_1 s_2} (A^T_0 A^T_1)_{s_0} (A^T_3 A^T_2)_{s_3} U_{s_0 s_1} U_{s_2 s_3} \left|0_{s_0} 0_{s_1} 0_{s_2} 0_{s_3}\right\rangle \tag{7.78}$$
$$= (A^T_0 A^T_1)_{s_0} (A^T_3 A^T_2)_{s_3} U^{\dagger}_{s_1 s_2} U_{s_0 s_1} U_{s_2 s_3} \left|0_{s_0} 0_{s_1} 0_{s_2} 0_{s_3}\right\rangle . \tag{7.79}$$

Next, we apply the entanglement swapping property on the set of qubit sets $\{s_0, s_1, s_2, s_3\}$, which leads to

$$\sqrt{p}(A^T_0 A^T_1)_{s_0} (A^T_3 A^T_2)_{s_3} U_{s_0 s_3} \left|0_{s_0} 0_{s_3}\right\rangle , \tag{7.80}$$

where

$$p = \begin{cases} \frac{1}{4^m}, & \text{if } U \text{ prepares } |\phi^+_n\rangle \\ \frac{1}{\binom{m}{k}^2}, & \text{if } U \text{ prepares } |\phi^m_k\rangle. \end{cases} \tag{7.81}$$

To prove QCSA, we assume as the induction hypothesis, that (7.75) is true. Taking $d \to d+2$ in (7.75) yields

$$\left(\prod_{k=1}^{d} \langle 0_{s_k}|\right)\left(\prod_{k=0}^{d/2-1} U^{\dagger}_{s_{2k+1},s_{2k+2}}\right)\left(\prod_{k=0}^{d/2}(A^T_{2k})_{s_{2k}}\right)$$
$$\times \left(\prod_{k=0}^{d/2}(A_{2k+1})_{s_{2k+1}}\right)\left(\prod_{k=0}^{d/2} U_{s_{2k},s_{2k+1}}\right)\left(\prod_{k=0}^{d+1}|0_{s_k}\rangle\right). \tag{7.82}$$

We now pull out of their respective products: the last two U terms $\left(U_{s_{d-2},s_{d-1}} \text{ and } U_{s_d,s_{d+1}}\right)$, the last four A terms $\left((A^T_{d-2})_{s_{d-2}}, (A_{d-1})_{s_{d-1}}, (A^T_d)_{s_d}, \text{and}(A_{d+1})_{s_{d+1}}\right)$, and the last $U^{\dagger}$ term $\left(U^{\dagger}_{s_{d-1},s_d}\right)$. This yields

$$\left(\prod_{k=1}^{d} \langle 0_{s_k}|\right)\left(\prod_{k=0}^{d/2-2} U^{\dagger}_{s_{2k+1},s_{2k+2}}\right)\left(\prod_{k=0}^{d/2-2}(A^T_{2k})_{s_{2k}}\right)\left(\prod_{k=0}^{d/2-2}(A_{2k+1})_{s_{2k+1}}\right)$$
$$\times \left(\prod_{k=0}^{d/2-1} U_{s_{2k},s_{2k+1}}\right)\Bigg[U^{\dagger}_{s_{d-1},s_d}(A^T_{d-2})_{s_{d-2}}(A_{d-1})_{s_{d-1}}(A^T_d)_{s_d}(A_{d+1})_{s_{d+1}} \tag{7.83}$$
$$\times U_{s_{d-2},s_{d-1}}U_{s_d,s_{d+1}}\left(\prod_{k=0}^{d+1}|0_{s_k}\rangle\right)\Bigg]. \tag{7.84}$$

Using the base case yields

$$\sqrt{p}\left(\prod_{k=1}^{d} \langle 0_{s_k}|\right)\left(\prod_{k=0}^{d/2-2} U^{\dagger}_{s_{2k+1},s_{2k+2}}\right)\left(\prod_{k=0}^{d/2-2}(A^T_{2k})_{s_{2k}}\right)\left(\prod_{k=0}^{d/2-2}(A_{2k+1})_{s_{2k+1}}\right)$$
$$\times \left(\prod_{k=0}^{d/2-1} U_{s_{2k},s_{2k+1}}\right)\left[\left(A^T_{d-2}A^T_{d-1}\right)_{s_{d-2}}(A_{d+1}A_d)_{s_{d+1}}U_{s_{d-2},s_{d+1}}\right]\left(\prod_{k=0}^{d+1}|0_{s_k}\rangle\right). \tag{7.85}$$

Absorbing the bracketed terms back into the products by relabeling $s_{d+1} \to s_d$, $A_{d+1}A_d \to A_{d-1}$, and $A^T_{d-2}A^T_{d-1} \to A^T_{d-2}$ yields

$$\sqrt{p}\left(\prod_{k=1}^{d} \langle 0_{s_k}|\right)\left(\prod_{k=0}^{d/2-2} U^{\dagger}_{s_{2k+1},s_{2k+2}}\right)\left(\prod_{k=0}^{d/2-1}(A^T_{2k})_{s_{2k}}\right)$$

$$\times \left(\prod_{k=0}^{d/2-1} (A_{2k+1})_{s_{2k+1}} \right) \left(\prod_{k=0}^{d/2-1} U_{s_{2k},s_{2k+1}} \right) \left(\prod_{k=0}^{d+1} |0_{s_k}\rangle \right). \tag{7.86}$$

Applying the induction hypothesis (7.75) yields

$$\sqrt{p^{d/2}} \left(\prod_{k=0}^{d-1} (A_k)_{s_{2d-1}} \right) U_{s_0,s_{2d-1}} |0\rangle , \tag{7.87}$$

which completes the proof. The final step is to measure the the first qubit set s_0. If the measurement succeeds (returns all zeros: $0_1 \ldots 0_m$) then the circuit is in the state

$$\sqrt{p^{(d+1)/2}} \left(\prod_{k=0}^{d-1} (A_k)_{s_{2d-1}} \right) U_{s_0,s_{2d-1}} |0\rangle . \tag{7.88}$$

because the probability of success is $\sqrt{p}$.

7.8 Scaling

In this section we discuss the scaling of various parameters involved in the QCSA algorithm. Let the initial circuit that one wishes to "squeeze" have m_i qubits and a depth of n_i. Let the final (squeezed) circuit have m_f qubits and a depth of n_f. Let d be an even integer between 1 and $m + 1$ and the factor by which the initial circuit is squeezed. The parameters of the final circuit (m_f and n_f) are given in terms of the parameters of the initial circuit (m_i and n_i) as follows:

$$m_f = dm_i, \tag{7.89}$$

$$n_f = O\left(\frac{n_i}{d}\right). \tag{7.90}$$

by the pigeon hole principle. The depth comparison equation (7.90) is written in big O notation because the true depth of the final circuit is $\lceil \frac{n_i}{d} \rceil$ plus $2d(U)$, twice the depth of U. If U prepares $|\phi_m^+\rangle$ then $d(U) = 2$. If U prepares $|\phi_k^m\rangle$ then $d(U) = O(m)$ if U is decomposed deterministically [9]. However, see the following chapter (8) for a potentially

shorter depth decomposition. In any case, for small m_i and large n_i, the term $\lceil \frac{n_i}{d} \rceil$ outweighs $2d(U)$. With these scalings (7.89 and 7.90) we can see that QCSA is most beneficial if one wishes to run a small-qubit, long-depth circuit on a large-qubit quantum computer that only allows short-depth circuits (and is substantially noisy). Fortunately, this is likely the direction that many NISQ era quantum computers are going [79].

We must also consider the scaling of the time it takes to run QCSA. Recall, that QCSA is a probabilistic algorithm, which means that it only succeeds with a certain probability. For the full or subspace version of QCSA (U prepares $|\phi_m^+\rangle$ or $|\phi_k^m\rangle$, respectively) the probabilities of success (p_{full} or p_{sub}, respectively) are

$$p_{\text{full}} = \frac{1}{2^{m_i(d-1)}} = O\left(\frac{1}{2^{m_i d}}\right), \tag{7.91}$$

$$p_{\text{sub}} = \frac{1}{\binom{m_i}{k}^{d-1}} = O\left(\frac{1}{m_i^{kd}}\right). \tag{7.92}$$

Note that this implies that the number of shots (runs of the quantum circuit) that are required for the results of the squeezed circuit to match the quality of the results of the original circuit scales inversely proportionally with p. That is, because the squeezed circuit is only equivalent to the original circuit (when run) with probability p, one must (on average) run the squeezed circuit p times to equal one run of the original circuit. Thus, the number of shots (s_{full} or s_{sub}) required for each version of QCSA (full or subspace, respectively) is given by

$$s_{\text{full}} = \frac{1}{p_{\text{full}}} = O\left(2^{m_i d}\right), \tag{7.93}$$

$$s_{\text{sub}} = \frac{1}{p_{\text{sub}}} = O\left(m_i^{kd}\right). \tag{7.94}$$

The time (t) it takes to run the experiment (run the quantum circuit s times) is given by the product of the time it takes to run the quantum circuit once (which scales as the depth n)

and the number of shots s. Thus, the experiment times for the initial and final circuits of full QCSA ($t_i^{(\text{full})}$ and $t_f^{(\text{full})}$, respectively) scale as

$$t_i^{(\text{full})} \propto n_i, \tag{7.95}$$

$$t_f^{(\text{full})} \propto s_{\text{full}} n_f = O\left(\frac{2^{m_i d}}{d}\right) n_i, \tag{7.96}$$

whereas the experiment times for the initial and final circuits of subspace QCSA (section 7.5), which are given by$t_i^{(\text{sub})}$ and ($t_f^{(\text{sub})}$ respectively, scale as

$$t_i^{(\text{sub})} \propto n_i, \tag{7.97}$$

$$t_f^{(\text{sub})} \propto s_{\text{sub}} n_f = O\left(\frac{m_i^{kd}}{d}\right) n_i. \tag{7.98}$$

Thus, the factor by which the experiment times increase can be expressed via

$$t_f^{(\text{full})} = O\left(\frac{2^{m_i d}}{d}\right) t_i^{(\text{full})}, \tag{7.99}$$

$$t_f^{(\text{sub})} = O\left(\frac{m_i^{kd}}{d}\right) t_i^{(\text{sub})}. \tag{7.100}$$

It can thus be seen that the subspace version of QCSA only takes polynomially longer to run as the number of qubits m increases, which is as compared to the full version of QCSA which takes exponentially longer to run. This is the main advantage of the subspace version of QCSA. The trade off is the longer circuit depth of U. Thus, the full version of QCSA should only be used for a very small number of qubits m. Note that both version's times grow exponentially with the squeezing factor d. Thus, although it would be wonderful to choose $d = n$ (the circuit depth) and squeeze our circuit to as short a depth as possible, this would not be advisable given the scaling of d. One potential sweet spot for d would be as a logarithm, that is

$$d_{\text{full}} = \frac{\log_2 n_i}{m_i}, \tag{7.101}$$

$$d_{\text{sub}} = \frac{\log_{m_i} n_i}{k}. \tag{7.102}$$

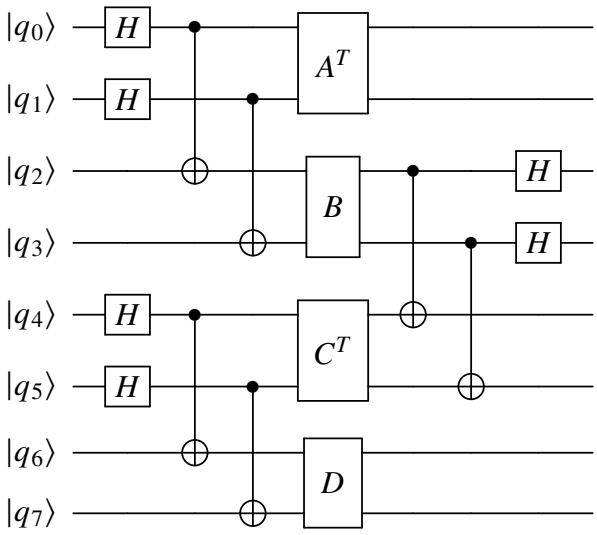
This choice of d for full QCSA would require the initial depth n_i to be exponential in the initial number of qubits m_i (that is, $n_i = 2^{cm_i}$) in order for $d = c$ (with $c > 1$). However, this choice of d for subspace QCSA would only require the initial depth n_i to be polynomial in the initial number of qubits m_i (that is, $n_i \propto m_i^{ck}$) to achieve the same. Plugging these values of d into 7.103 and 7.104, respectively, yields

$$t_f^{(\text{full})} = O\left(\frac{n_i}{\log_2 n_i}\right) t_i^{(\text{full})}, \tag{7.103}$$

$$t_f^{(\text{sub})} = O\left(\frac{n_i}{\log_{m_i} n_i}\right) t_i^{(\text{sub})}, \tag{7.104}$$

which is sub-linear scaling. All of this again reiterates that QCSA is best suited for circuits are much deeper than they are tall (number of qubits).

Finally, we discuss the qubit-connectivity required to run QCSA. Assuming that all gate columns A only require linear connectivity, then the full version of QCSA only requires grid connectivity. For example, the following squeezed quantum circuit for full QCSA with $(m_i, n_i, d) = (2, 4, 2)$:



(7.105)

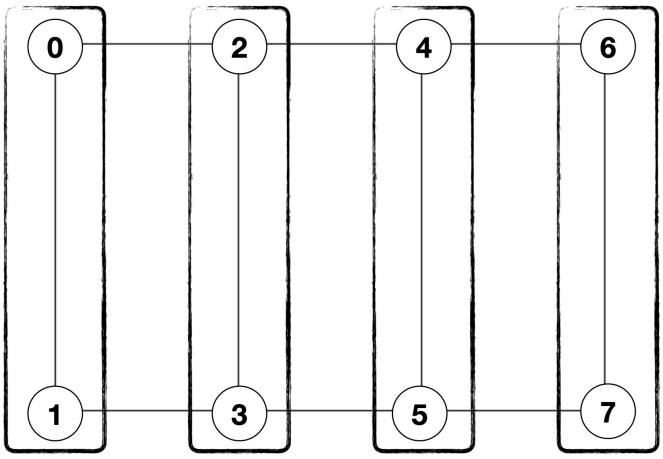


Figure 7.1: Graph of grid qubit-connectivity. Circled numbers represent qubits, black lines represent that two qubits are connected, and black rectangles represent gates A.

can be run on a quantum computer with the grid architecture found on Figure 7.1, where the black rectangles represent the two-qubit gates (A^T, B, C^T, and D). They inscribe the qubits upon which their corresponding two-qubit gate acts. In the case of subspace QSCA, as long as U can prepare $\left|\phi_k^n\right\rangle$ with linear-connectivity (reference [9] and chapter 8), then the algorithm itself only requires linear connectivity 5.4.

7.9 Demonstration

Here we given a demonstration of QCSA. We have developed code that performs QCSA: taking in an arbitrary quantum circuit and transforming it to the squeezed version of itself. We tested the algorithm using the following procedure: for various initial circuit depths (n), we estimated the density matrix (ρ_{est}) of the final state of each circuit (original and squeezed) using quantum state tomography (6.56) and compared its corresponding exact density matrix (ρ_{ext}) by taking the fidelity of the two. The definition of fidelity that we used can be defined as follows. Let ρ and σ be two density matrices, where the density matrix ρ of a state $|\phi\rangle$ is defined as ($\rho = |\psi\rangle \langle\psi|$). Then their fidelity $F(\sigma, \rho)$ is defined as

the trace of their product

$$\mathrm{Tr}(\sigma\rho). \tag{7.106}$$

Quantum state tomography is the process by which the density matrix of a n-qubit state $|\psi\rangle$ is estimated as

$$\rho = \frac{1}{2^n} \sum_{\alpha_1,\ldots,\alpha_n=0}^{3} S_{\alpha_1,\cdots,\alpha_n} \sigma^{(\alpha_0)} \otimes \ldots \otimes \sigma^{(\alpha_n)}, \tag{7.107}$$

where $\sigma^0 = I$, $\sigma^1 = X$, $\sigma^2 = Y$, and $\sigma^3 = Z$ are the Pauli operators and

$$S_{\alpha_1,\ldots,\alpha_n} = \langle\psi|\sigma^{(\alpha_0)} \otimes \ldots \otimes \sigma^{(\alpha_n)}|\psi\rangle\,, \tag{7.108}$$

are the expectation values of the Pauli-strings $\sigma^{(\alpha_0)} \otimes \ldots \otimes \sigma^{(\alpha_n)}$ which are estimated using the method described in subsection 3.4.2. We studied the case with three initial qubits ($m_i = 3$), a squeezing factor of $d = 2$, and let the initial depth run over $n_i = 2, 3, \ldots, 18$. The original circuits were populated with random three-qubit gates and each estimation of fidelity was averaged over five runs. The number of shots used for the squeezed circuit s_f was $s_f = 2^{m(d-1)} s_i = 8 s_i$ where $s_i = 2^{11}$ is the number of shots used for the original circuit. The results were obtained from a noisy simulation of the quantum circuits on a classical computer. It can be seen in Figure 7.2 that the the fidelity estimated from the squeezed circuit is always greater than that estimated from the original circuit, the difference becoming more pronounced as the absolute difference in their depths grows with the increase in the original circuits depth n_i. As the squeezing factor was $d = 2$, that means the ratio of the circuit depths n_f/n_i approaches 2 as n_i grows. Here, n_f is the squeezed circuit depth.

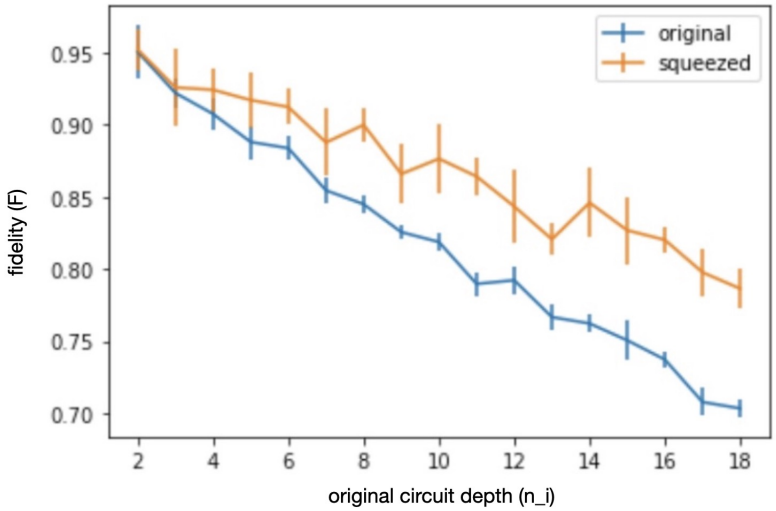


Figure 7.2: Comparison of fidelities of estimated and exact density matrices of both original and squeezed circuits.

7.10 Conclusion

In this chapter, we have presented the quantum circuit squeezing algorithm (QCSA), a novel quantum algorithm that trades more qubits and shots for a shorter circuit depth. To explain the algorithm, we first introduced the concepts of maximal entanglement, the ricochet property, and the entanglement swapping property. It was then shown how the entanglement property can be extended recursively. Additionally, a modification of QCSA to deal with Hamming weight preserving subspaces was introduced which ultimately improved the time scaling qubit connectivity requirements of the algorithm. It was then shown how all of these techniques could be combined to create QCSA. An analysis of the scaling of various complexities of the algorithm was conducted. Finally, we demonstrated the algorithm by running noisy simulations of it to show that the resulting squeezed circuits were less affected by the noise due to their shorter depth. As NISQ era devices require short depth circuits due to their high noise levels and are rapidly growing in their number

of qubits, QCSA has the potential to be very beneficial in the near future by allowing researchers to run quantum circuits that otherwise would have been lost to the noise of our era's quantum machines.

CHAPTER 8

VARIATIONAL PREPARATION OF DICKE STATES

8.1 Introduction

Several chapters of this thesis involve the use of Dicke states. Chapter 4 used them to construct the ansatz for one mapping of the problem and again used them to map the problem to a smaller set of qubits. Additionally, chapter 5 used them as an initialization strategy for a novel type of ansatz to solve a model in an extreme case. Finally, chapter 7 relied on the construction of them for a modification of the algorithm for Hamming weight preserving gates. Therefore, it is clear that an efficient algorithm to construct Dicke states would be very beneficial to many areas of the application of NISQ era algorithms to many-body nuclear physics. While a deterministic method exists to do so that is linear in the number of qubits ([9] and [2]) the quantum circuit constructed is often still too long to be implemented on noisy quantum devices of the NISQ era. As an alternative we present here a short-depth, variational algorithm with the potential to prepare Dicke states with a shorter depth than the previously mentioned algorithm.

First, let us start with a few definitions. Recall that a Dicke state $|D_k^n\rangle$ is the equal superposition of all n-qubit states $|x\rangle$ with Hamming weight $\mathrm{wt}(x) = k$; that is

$$\left|D_k^n\right\rangle = \binom{n}{k}^{-\frac{1}{2}} \sum_{\substack{x\in\{0,1\}^n \\ \mathrm{wt}(x)=k}} |x\rangle \,. \tag{8.1}$$

For example,

$$\left|D_2^4\right\rangle = \frac{1}{\sqrt{6}} \left(|1100\rangle + |1010\rangle + |1001\rangle + |0110\rangle + |0101\rangle + |0011\rangle \,.\right) \tag{8.2}$$

A Dicke-like state $\left|Dl_k^n(m)\right\rangle$ is the equal superposition of m, n-qubit states $|x\rangle$ with Hamming weight $\mathrm{wt}(x) = k$ up to possible relative phases of +1,

$$\left|Dl_k^n\right\rangle = \binom{n}{m}^{-\frac{1}{2}} \sum_{\substack{x\in\{0,1\}^n \\ \mathrm{wt}(x)=k}} (-1)^{f(x)} g(x) \left|x\right\rangle , \tag{8.3}$$

where $g(x) = 0, 1$ such that $g(x) = 1$ only for m states x. Thus a Dicke-like state is a phased Dicke state that's missing some bit-string states in its superposition. An example would be

$$\left|Dl_2^4(4)\right\rangle = \frac{1}{\sqrt{6}} \left(|1100\rangle - |1010\rangle + |1001\rangle - |0110\rangle \, .\right) \tag{8.4}$$

As discussed in the motivation section, we seek to find a quantum circuit that prepares the Dicke states with the shortest circuit depth possible. Here I propose a hybrid quantum-classical algorithm called the Variational State Preparation Algorithm (VSPA) to find such a circuit. VSPA is hybrid because it uses both a quantum and classical computer. The quantum computer's role is to implement a parameterized quantum circuit that tries to implement a Dicke state. Upon measurement of the quantum computer, a classical computer does post-measurement calculations to calculate a cost function which characterizes the overlap of the prepared state with the desired Dicke state. A classical optimization algorithm varies the parameters of the quantum circuit in order to minimize the cost function as much as possible. This results in the quantum circuit preparing a quantum state that has a desired overlap with the desired Dicke state.

8.2 The Algorithm

The variational quantum circuit is parameterized by a set of parameters that are tuned by the classical optimization algorithm. The circuit consists of two parts: alternating k-state

preparation and variational mixing. First, the quantum computer is put into an alternating k-state which is a particular quantum state with Hamming weight k. For n qubits, the alternating k-state of type t = I, II, III, $|A_k^n\rangle_t$ is the quantum state whose qubits alternate between 1 and 0 for the middle $2k$ qubits and are all 0 for the rest,

$$\left|A_k^n\right\rangle_{\mathrm{I}} = |...f(i)...\rangle \mid \begin{cases} f(i) = i+1 \pmod 2, & \left\lfloor \frac{n}{2} - 1 \right\rfloor - (k-1) \leq i \leq \left\lfloor \frac{n}{2} - 1 \right\rfloor + (k-1) \\ f(i) = 0, & \text{otherwise} \end{cases} \tag{8.5}$$

$$\left|A_k^n\right\rangle_{\mathrm{II}} = |...f(i)...\rangle \mid \begin{cases} f(i) = i \pmod 2, & \left\lfloor \frac{n}{2} - 1 \right\rfloor - (k-1) \leq i \leq \left\lfloor \frac{n}{2} - 1 \right\rfloor + (k-1) \\ f(i) = 0, & \text{otherwise} \end{cases} \tag{8.6}$$

$$\left|A_k^n\right\rangle_{\mathrm{III}} = \frac{1}{\sqrt{2}} \left(\left|A_k^n\right\rangle_{\mathrm{I}} + \left|A_k^n\right\rangle_{\mathrm{II}} \right), \tag{8.7}$$

where $k = 1, ..., \lfloor k/2 \rfloor$. For example

$$\left|A_2^6\right\rangle_{\mathrm{I}} = |010100\rangle\,, \tag{8.8}$$

$$\left|A_2^6\right\rangle_{\mathrm{II}} = |001010\rangle\,, \tag{8.9}$$

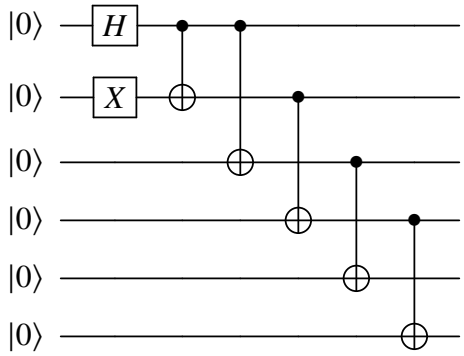
$$\left|A_2^6\right\rangle_{\mathrm{III}} = |0\rangle \frac{|1010\rangle + |0101\rangle}{\sqrt{2}} |0\rangle\,. \tag{8.10}$$

To prepare $|A_k^n\rangle_t$ for $k = \lfloor k/2 \rfloor + 1, ..., n$ one would finish by applying X gates to all qubits since

$$\left|A_{n-k}^n\right\rangle_t = X^{\otimes n} \left|A_k^n\right\rangle_t\,. \tag{8.11}$$

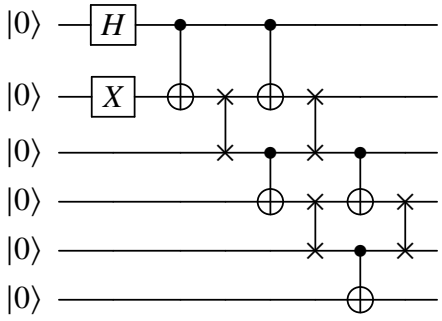
For n qubits, the alternating k-state of type I (II) can be prepared on a quantum computer by applying an X gates to all qubits that must be 1 (a quantum circuit of depth one). The

alternating k-state of type III can be prepared with a circuit of depth 3 with grid connectivity and depth $O(n)$ with linear connectivity: Given a grid architecture, the n-qubit alternating k-state of type III can be implemented with the following constant (2) depth circuit. For example, the circuit for $\left|A_2^6\right\rangle$ is shown below for six qubits:



(8.12)

The first two rows prepare the Bell state $|\psi^+\rangle = (|01\rangle + |10\rangle)/\sqrt{2}$ and each subsequent pair of CNOTs "copies" the state of each term of the Bell state and adds it onto that term. Given a linear nearest-neighbor (LNN) architecture the n-qubit alternating k-state of type III can be implemented with the following $2n + 1$ depth circuit. For example, the circuit for $\left|A_2^6\right\rangle$ is shown below for six qubits:



(8.13)

Once an alternating k-state has been prepared, the next step is to apply layers of parameterized, two-qubit, Hamming weight preserving gates called partial-SWAP gates.

The partial-SWAP gate parameterized by θ is defined to be

$$\text{pSWAP}(\theta) = e^{i\theta(XY-YX)/2} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{8.14}$$

Being in the group $SO(4)$, the pSWAP gate can be decomposed into a circuit of depth at most five, containing at most two CNOTs and twelve single-qubit rotations gates [4]. Note that pSWAP(θ) corresponds to a rotation in the subspace $\{|01\rangle, |10\rangle$ by the angle θ (leaving the subspace $\{|00\rangle, |11\rangle$ unchanged):

$$\text{pSWAP}(\theta)\,|00\rangle = |00\rangle \tag{8.15}$$

$$\text{pSWAP}(\theta)\,|01\rangle = \cos\theta\,|01\rangle + \sin\theta\,|10\rangle \tag{8.16}$$

$$\text{pSWAP}(\theta)\,|10\rangle = \cos\theta\,|10\rangle - \sin\theta\,|01\rangle \tag{8.17}$$

$$\text{pSWAP}(\theta)\,|11\rangle = |11\rangle, \tag{8.18}$$

thus preserving Hamming weight. Using pSWAP gates as mixing gates will always result in a Dicke-like state since there are no complex phases in their matrix representation (only a minus sign). A mixing layer is defined to be a single column of parameterized pSWAP gates being applied in parallel. The variational mixing part of the VSPA algorithm consists of several mixing layers. The first layer is applied to the middle $2k$ qubits. Each subsequent layer is applied to two additional qubits than the layer before - the qubit directly above and the qubit directly below the previous set of qubits (unless such qubit does not exist). The pseudo-code for the variational mixing part is given below:

Algorithm 8.1 Variational Mixing Algorithm

Input: Number of qubits n, Hamming weight k, number of layers l, set of angles $\{\theta\}$.
Output: Quantum circuit implementation of variational mixing section of VSPA.

```
index_list = [ ]

i = ⌊n/2 − 1⌋ − (k − 1)
while i ≤ ⌊n/2 − 1⌋ + (k − 1) do
    index_list.append(i)
    i+ = 2
end while

theta_index = 0
for 0 ≤ l ≤ layer do
    if l ≠ 0 then new_index_list = [ ]
        for i in index_list do
            new_index_list.append(i − 1)
        end for
        new_index_list.append(i + 1)
        index_list=new_index_list
    end if
    for q in index_list do
        if 0 ≤ q < n − 1 then
            Append pSWAP(θ[theta_index]) to qubits q and q+1 of the quantum circuit
            theta_index + = 1
        end if
    end for
end for
```

8.3 Calculating a Cost Function

As discussed above, our use of pSWAP gates as mixing gates will always result in a state that is the superposition of terms that all have the same Hamming weight. This should make it easier to minimize a cost function that measures the overlap of the variationally prepared state and a desired Dicke-like state. One good cost-function to quantify how close

a variationally prepared state is to an n-qubit, m-term Dicke-like state $\left|Dl_k^n(m)\right\rangle$ would be:

$$f_{\text{cost}}(|\psi\rangle) = \text{Var}\{|\langle x|\psi\rangle|^2 | x \in \{0,1\}^n | \text{wt}(x) = k\}, \tag{8.19}$$

That is, the variance of the list of probabilities of measuring every possible $|x\rangle$ with Hamming weight k. The variance makes a good cost function as it measures the deviations of a set from their average and the set of probabilities of an exact Dicke state would all be exactly zero, thus having a variance of zero. The closer the probabilities are to their average, the smaller the variance will be. These probabilities can be estimated by preparing and measuring the VSPA circuit multiple times and dividing the number of times one measured $|x\rangle$ by the total number of measurements made. The variance of the coefficients' absolute value squared is minimized when they are all equal (equal super-position) and is a measure of how close they are to being equal. Now, there are two variables to play with here that one may like to control: $g_n = 1 - \sum_x g(x)$ and $f_n = 1 - \sum_x f(x)$. Minimizing the first means maximizing the number of terms in the Dicke-like state, making the state a closer approximation to a true Dicke state. Minimizing the second means minimizing the number of terms that have a minus sign in front of them, again making the state a closer approximation to a true Dicke state. Another good candidate for a cost-function would be

$$f_{\text{cost}}(|\psi\rangle) = 1 - \frac{1}{|x(l')|}\left(\sum_{x \in x(l')} \sqrt{|\langle x|\psi\rangle|^2}\right)^2, \tag{8.20}$$

where $l' \leq l$, which is one minus the square of the sum of the square roots of the probabilities. It is equal to one minus the overlap between our prepared state $|\psi\rangle$ and a desired non-phased Dicke-like state if one assumes no phases in the prepared state. Even if the prepared state does have phases, it should still serve as a good cost-function because it minimizes the difference between the absolute squares of the coefficients. Note

that f_n cannot be controlled using either of these cost function alone because using the probabilities (absolute values squared) of the coefficients erases all plus-or-minus phase information. However, it may be discovered that having plus-or-minus phased Dicke states is still beneficial in some circumstances such as in the multi-configuration method (subsection 5.3.7 of chapter 5) as the solution to the pairing model for large g is not exactly a Dicke state but one close to it anyways.

8.4 Results

First, we created code to deterministically prepare Dicke states using the method given in [9]. We the created code to run my novel VSPA algorithm using either of the first two cost-functions discussed above. We have used this algorithm to prepare Dicke state approximations to $|D_2^n\rangle$ (where $n = 4, 5, 6$) that each have a higher overlap with the true Dicke state than their deterministically ([9]) prepared counterparts when simulated with noise (Figure 8.1). This is because, as seen in Figure 8.2, while the deterministic preparation would in theory, create a perfect Dicke state, its depth is so much longer than our variational circuit that, when noise is included, creates a worse approximation to the true Dicke state. The idea is that, even though the variational method may not be able to exactly prepare an exact Dicke state (which would require exactly minimizing the cost function in a noisy environment), it has the potential to get closer than the deterministic method once one accounts for the extra noise that befalls the latter. In the bottom half of Figure 8.2, we have a histogram of the measurements obtained through noisy simulation. The measurements that have a Hamming weight of two (the only ones which should have been measured) are highlighted in red. All non-feasible (with a Hamming weight other than two) are highlighted in blue. With this, it can clearly be seen that in this case, the

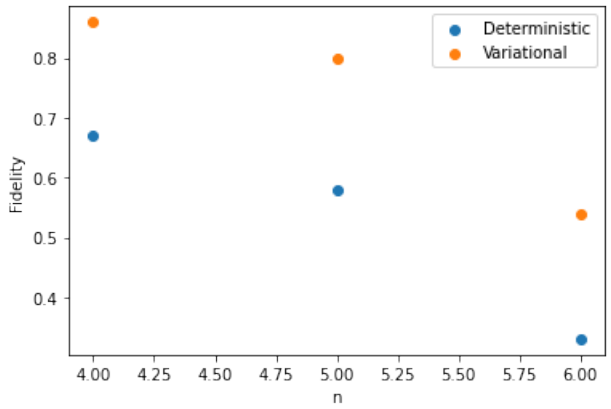


Figure 8.1: Overlap of deterministically and variationally prepared Dicke-states with true Dicke-state.

variational method is much less affected by noise as the non-feasible states are measured with much less frequency than in the deterministic method. Next, the number of CNOTs is noted to be significantly less for the variational method, likely due to the large number of the three-qubit, double controlled gates that had to be decomposed into several CNOTs each for the deterministic method. This was also a likely factor in improving the results as the CNOT is noisier than the single-qubit gates. Finally, we note the high number of CNOTs in the the deterministic method circuit which connect non-neighboring qubits. As we simulated the results on a noisy model that assumed only linear connectivity of qubits, the compiler for the deterministic method's circuit had to add several more swap gates (which require 3 CNOTs) compared to the variational method's in order to move the non-neighboring qubits next to one another. This is because the deterministic method's circuit has six non-neighboring CNOTs, compared to the variational method's two.

8.5 Conclusion

In this section, we developed the variational Dicke state preparation algorithm (VSPD), a novel variational algorithm to prepare Dicke and Dicke-like states. We laid forth a short-

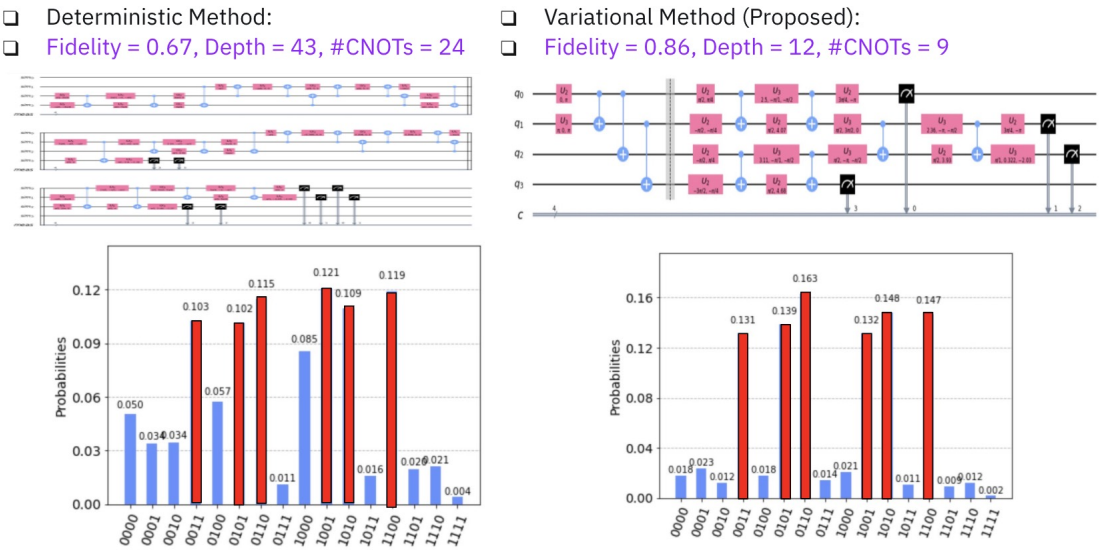


Figure 8.2: Comparison of deterministic and variational methods to prepare the Dicke state $|D_2^4\rangle$.

depth variational ansatz which can efficiently search the appropriate Hamming weight restricted subspace. We also discussed different cost functions which could be used for the minimization subroutine. Finally, we presented results obtained from testing the algorithm with noisy simulations and bench-marked them favorably against a previously developed deterministic method [9]. In the future, one might wish to compare and contrast different ansatz and cost functions to prepare Dicke state $|D_k^n\rangle$ for even more values of n and k. Additionally, one could look into determining a way to fix the phases $f(x)$ of the terms of the variational Dicke state prepared. One potential way to do this would simply to have the ansatz and cost-function be identical to that introduced for the pairing model in chapter 5) with the pairing strength g set to zero. This is because, as proven in subsection 5.3.7, the Dicke states are the eigenstates of said model. We conclude this section by reiterating the widespread potential use of Dicke states: both as part of ansatz construction for many-body nuclear models, and as part of the entangling process of the subspace QCSA algorithm.

CHAPTER 9

CONCLUSION

In this thesis, we have provided the basis for which future quantum algorithms can be developed to solve many-body nuclear problems. For the Lipkin model, we developed techniques to make the ansatz for different mappings of the problem lower-depth. We also developed a mapping that cut the number of required qubits in half. For the pairing model, we estimated the ground state energy with the variational quantum eigensolver (VQE) and bench-marked the results against the classical algorithm pair coupled cluster theory (pCCD). We also developed techniques to reduce the circuit depth for the unitary pair coupled cluster doubles (UpCCD) ansatz for quantum computers with circular connectivity that has a lower depth than the ansatz for devices with linear connectivity (5.4). We also developed an algorithm for finding the energies of excited states for the model. Finally, we developed a novel ansatz in which one starts the quantum computer in a Dicke state in order to solve the pairing model for large values constant pairing strength g. In the collective neutrino oscillations chapter, we presented the first digital simulation of the model on a real quantum computer. We used this to study the time-evolution and entanglement of the system. We applied error mitigation techniques to improve the accuracy of our results. Then we presented the quantum circuit squeezing algorithm and showed how it could be used to reduce the circuit depth and therefore improve the results of arbitrary quantum circuits. Finally, we presented our novel algorithm to variationally prepared Dicke states.

As this work was meant to serve as a springboard of which future developments in the field could be accomplished, we discuss here several future extensions and applications of this work. For the Lipkin model, one could explicitly compare the different mappings and

ansatzes discussed to have better evidence that the one we choose to pursue is truly the best choice. One could do this by running tests on the gradients of the wave-function to see if either of the methods lends itself more easily to the problem of the barren plateau (exponentially vanishing gradients). In terms of the pairing model, one would desire to compare our iterative quantum excited states algorithm to other algorithms that attempt to accomplish the same thing. Additionally, one would like to run the VQE algorithms of this section on the quantum computers of the near future (via the cloud) that would allow for the frequent calls to the quantum computer that is required for VQE. As for the collective neutrino oscillation simulations, one would like to compare various permutations of the Trotter terms that make up the ansatz to see if putting the terms with the highest commutator with the Hamiltonian first would improve the results on an actual quantum computer. The reasoning behind this is that, because the qubits decohere exponentially, putting the most important terms of the ansatz first would give said terms the least noisy effect on the quantum state. For the quantum circuit squeezing algorithm, one could apply it to a pre-existing algorithm like VQE to see if it improves the accuracy of the results. Finally, the variational preparation of Dicke states could similarly be applied to pre-existing algorithms that would benefit from the initialization of their circuits into a Dicke state. In the far future, one would like to see the improved quantum computers of that time be used to tackle the problems of nuclear physics that are too complex for today's classical computers, serving as a light in the deep hole of our ignorance, shone towards the hidden treasures buried deep below.

BIBLIOGRAPHY

[1] "A restricted class of exact eigenstates of the pairing-force Hamiltonian". In: *Phys. Lett.* 3.6 (1963), pp. 277–279. DOI: 10.1016/0031-9163(63)90259-2.

[2] Shamminuj Aktar et al. "A Divide-and-Conquer Approach to Dicke State Preparation". In: *IEEE TQE* 3 (2022), pp. 1–16. DOI: 10.1109/TQE.2022.3174547.

[3] Héctor Abraham et al. *Qiskit: An Open-source Framework for Quantum Computing*. 2019. DOI: 10.5281/zenodo.2562110.

[4] F. T. Arecchi et al. "Atomic Coherent States in Quantum Optics". In: *Phys. Rev. A* 6 (6 1972), pp. 2211–2237. DOI: 10.1103/PhysRevA.6.2211.

[5] A. B. Balantekin and Y. Pehlivan. "Solutions of nuclear pairing". In: *Phys. Rev. C* 76 (5 2007), p. 051001. DOI: 10.1103/PhysRevC.76.051001.

[6] K. Banaszek et al. "Maximum-likelihood estimation of the density matrix". In: *Phys. Rev. A* 61 (1 1999), p. 010304. DOI: 10.1103/PhysRevA.61.010304.

[7] J. Bardeen, L. N. Cooper, and J. R. Schrieffer. "Theory of Superconductivity". In: *Phys. Rev.* 108 (5 1957), pp. 1175–1204. DOI: 10.1103/PhysRev.108.1175.

[8] Rodney J. Bartlett, Stanislaw A. Kucharski, and Jozef Noga. "Alternative coupled-cluster ansätze II. The unitary coupled-cluster method". In: *Chem. Phys. Lett.* 155.1 (1989), pp. 133–140. DOI: 10.1016/S0009-2614(89)87372-5.

[9] Andreas Bartschi and Stephan Eidenbenz. "Deterministic Preparation of Dicke States". In: *Phys. Rev. D* 104 (6 2021), p. 063009. DOI: 10.1103/PhysRevD.104.063009.

[10] Nicole F. Bell, Andrew A. Rawlinson, and R.F. Sawyer. "Speed-up through entanglement—many-body effects in neutrino processes". In: *Phys. Lett. B* 573 (2003), pp. 86–93. DOI: 10.1016/j.physletb.2003.08.035.

[11] Jakob Bernoulli. *Ars Conjectandi: Usum amp; Applicationem Praecedentis*

Doctrinae in Civilibus, Moralibus amp; Oeconomicis.

[12] A. Bohr, B. R. Mottelson, and D. Pines. "Possible Analogy between the Excitation Spectra of Nuclei and Those of the Superconducting Metallic State". In: *Phys. Rev.* 110 (4 1958), pp. 936–938. DOI: 10.1103/PhysRev.110.936.

[13] Zhenyu Cai. *Multi-exponential Error Extrapolation and Combining Error Mitigation Techniques for NISQ Applications*. 2020. arXiv: 2007.01265 [quant-ph].

[14] J. Carlson et al. "Quantum Monte Carlo methods for Nucl. Phys." In: *Rev. Mod. Phys.* 87 (3 2015), pp. 1067–1118. DOI: 10.1103/RevModPhys.87.1067.

[15] D. M. Ceperley. "Path integrals in the theory of condensed helium". In: *Rev. Mod. Phys.* 67 (2 1995), pp. 279–355. DOI: 10.1103/RevModPhys.67.279.

[16] Michael J. Cervia et al. "Entanglement and collective flavor oscillations in a dense neutrino gas". In: *Phys. Rev. D* 100 (8 2019), p. 083001. DOI: 10.1103/PhysRevD.100.083001.

[17] Michael J. Cervia et al. "Lipkin model on a quantum computer". In: *Phys. Rev. C* 104 (2 2021), p. 024305. DOI: 10.1103/PhysRevC.104.024305.

[18] Sovan Chakraborty et al. "Collective neutrino flavor conversion: Recent developments". In: *Nucl. Phys. B* 908 (2016). Neutrino Oscillations: Celebrating the Nobel Prize in Physics 2015, pp. 366–381. DOI: 10.1016/j.nuclphysb.2016.02.012.

[19] Jerry Chow, Oliver Dial, and Jay Gambetta. *IBM quantum breaks the 100-qubit processor barrier*. Feb. 2022. URL: https://research.ibm.com/blog/127-qubit-quantum-processor-eagle.

[20] Wikimedia Commons. *File:Bloch sphere.svg — Wikimedia Commons, the free media repository*. 2022. URL: https://commons.wikimedia.org/w/index.php?title=File:Bloch_sphere.svg&oldid=630223048.

[21] Jakub Czartowski et al. “Isoentangled Mutually Unbiased Bases, Symmetric Quantum Measurements, and Mixed-State Designs”.
In: *Phys. Rev. Lett.* 124 (9 2020), p. 090503.
DOI: 10.1103/PhysRevLett.124.090503.

[22] D. J. Dean and M. Hjorth-Jensen.
“Pairing in nuclear systems: from neutron stars to finite nuclei”.
In: *RMP* 75.2 (2003), pp. 607–656. DOI: 10.1103/revmodphys.75.607.

[23] R. H. Dicke. “Coherence in Spontaneous Radiation Processes”.
In: *Phys. Rev.* 93 (1 1954), pp. 99–110. DOI: 10.1103/PhysRev.93.99.
URL: https://link.aps.org/doi/10.1103/PhysRev.93.99.

[24] P. Dirac. “The quantum theory of the emission and absorption of radiation”.
In: *Proc. R. Soc.* 144 (767 1927), pp. 243–265. DOI: 10.1098/rspa.1927.0039.

[25] Huaiyu Duan, George M. Fuller, and Yong-Zhong Qian.
“Collective Neutrino Oscillations”.
In: *Annu. Rev. Nucl. Part. Sci.* 60.1 (2010), pp. 569–594.
DOI: 10.1146/annurev.nucl.012809.104524.

[26] Huaiyu Duan et al. “Simulation of coherent nonlinear neutrino flavor transformation in the supernova environment: Correlated neutrino trajectories”.
In: *Phys. Rev. D* 74 (10 2006), p. 105014. DOI: 10.1103/PhysRevD.74.105014.

[27] J. Dukelsky et al. “Exactly solvable pairing Hamiltonian for heavy nuclei”.
In: *Phys. Rev. C* 84 (6 2011), p. 061301. DOI: 10.1103/PhysRevC.84.061301.

[28] E. F. Dumitrescu et al. “Cloud Quantum Computing of an Atomic Nucleus”.
In: *Phys. Rev. Lett.* 120 (21 2018), p. 210501.
DOI: 10.1103/PhysRevLett.120.210501.

[29] E. F. Dumitrescu et al. “Cloud Quantum Computing of an Atomic Nucleus”.
In: *Phys. Rev. Lett.* 120 (21 2018), p. 210501.
DOI: 10.1103/PhysRevLett.120.210501.

[30] Vincent E. Elfving et al. “Simulating quantum chemistry in the seniority-zero space on qubit-based quantum computers”.
In: *Phys. Rev. A* 103 (3 2021), p. 032605. DOI: 10.1103/PhysRevA.103.032605.

[31] Suguru Endo, Simon C. Benjamin, and Ying Li.
"Practical Quantum Error Mitigation for Near-Future Applications".
In: *Phys. Rev. X* 8 (3 2018), p. 031027. DOI: 10.1103/PhysRevX.8.031027.

[32] Coester F. "Bound states of a many-particle system".
In: *Nucl. Phys.* 7 (1958), pp. 421–424. DOI: 10.1016/0029-5582(58)90280-3.

[33] Richard P. Feynman. "Simulating physics with computers".
In: *Int. J. Theor. Phys.* 21 (1982), pp. 467–488. DOI: 10.1007/BF02650179.

[34] V. Fock. "Konfigurationsraum und zweite Quantelung".
In: *Zeitschrift fur Physik* 75.9-10 (1932), pp. 622–647. DOI: 10.1007/BF01344458.

[35] Alexander Friedland and Cecilia Lunardini.
"Do many-particle neutrino interactions cause a novel coherent effect?"
In: *JHEP* 2003.10 (2003), pp. 043–043. DOI: 10.1088/1126-6708/2003/10/043.
URL: 10.1088/1126-6708/2003/10/043.

[36] Juan Carlos Garcia-Escartin and Pedro Chamorro-Posada.
"swap test and Hong-Ou-Mandel effect are equivalent".
In: *Phys. Rev. A* 87 (5 2013), p. 052330. DOI: 10.1103/PhysRevA.87.052330.

[37] T. Giurgica-Tiron et al.
"Digital zero noise extrapolation for quantum error mitigation". In: *2020 IEEE International Conference on Quantum Computing and Engineering (QCE)*. 2020, pp. 306–316. DOI: 10.1109/QCE49297.2020.00045.

[38] M. C. Gonzalez-Garcia and Yosef Nir.
"Neutrino masses and mixing: evidence and implications".
In: *Rev. Mod. Phys.* 75 (2 2003), pp. 345–402.
DOI: 10.1103/RevModPhys.75.345.

[39] Benjamin Hall et al.
"Simulation of collective neutrino oscillations on a quantum computer".
In: *Phys. Rev. D* 104 (6 2021), p. 063009. DOI: 10.1103/PhysRevD.104.063009.

[40] W. Hartree. "Self-consistent field, with exchange, for beryllium".
In: *Proc. R. Soc.* 50 (869 1935), pp. 9–33. DOI: 10.1098/rspa.1935.0085.

[41] Andre He et al. “Zero-noise extrapolation for quantum-gate error mitigation with identity insertions”. In: *Phys. Rev. A* 102 (1 2020), p. 012426. DOI: 10.1103/PhysRevA.102.012426.

[42] Thomas Henderson et al. “Seniority-based coupled cluster theory”. In: *J. Chem. Phys.* 141 (2014). DOI: 10.1063/1.4904384.

[43] Scott Hill and William K. Wootters. “Entanglement of a Pair of Quantum Bits”. In: *Phys. Rev. Lett.* 78 (26 1997), pp. 5022–5025. DOI: 10.1103/PhysRevLett.78.5022.

[44] Morten Hjorth-Jensen. URL: http://nucleartalent.github.io/Course2ManyBodyMethods/doc/pub/secondquant/html/secondquant-bs.html.

[45] 5qubit backed: IBM Q team. *IBM Vigo backend specification V1.0.2*. Retrieved from https://quantum-computing.ibm.com. 2020. URL: https://quantum-computing.ibm.com.

[46] Rodolfo Id Betan. “Richardson’s Solutions in the Real– and Complex–Energy Spectrum”. In: *J. Phys. Conf. Ser.* 839 (2017), p. 012003. DOI: 10.1088/1742-6596/839/1/012003.

[47] Piet Van Isacker and Kristiaan Heyde. *Exactly solvable models of nuclei*. 2014. URL: http://www.scholarpedia.org/article/Exactly_solvable_models_of_nuclei#Pan98.

[48] P. Jordan and E. Wigner. “Über das Paulische Äquivalenzverbot”. In: *Zeitschrift für Physik* 47 (1928). DOI: 10.1007/BF01331938.

[49] Abhinav Kandala et al. “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets”. In: *Nature* 549 (2017), pp. 242–246. DOI: 10.1038/nature23879.

[50] Alexei Y. Kitaev. “Quantum measurements and the Abelian Stabilizer Problem”. In: *Electron. Colloquium Comput. Complex.* 3 (1996).

[51] Ian D. Kivlichan et al. “Quantum Simulation of Electronic Structure with Linear Depth and Connectivity”. In: *Phys. Rev. Lett.* 120 (11 2018), p. 110501.

DOI: 10.1103/PhysRevLett.120.110501.

[52] J.G. Lietz et al. “Computational Nucl. Phys. and Post Hartree-Fock Methods”. In: *An advanced course in computational Nucl. Phys. bridging the scales from quarks to neutron stars*. Springer International Publishing, 2017, pp. 327–333.

[53] J.G. Lietz et al. “Computational Nucl. Phys. and Post Hartree-Fock Methods”. In: *An advanced course in computational Nucl. Phys. bridging the scales from quarks to neutron stars*. Springer International Publishing, 2017, pp. 332–333.

[54] Seth Lloyd. “Universal Quantum Simulators”.
In: *Science* 273.5278 (1996), pp. 1073–1078.
DOI: 10.1126/science.273.5278.1073.

[55] J. Lopes dos Santos. URL: https://www.fc.up.pt/pessoas/jlsantos/wp-content/myuploads/NewperspectiveHartreeFock.pdf.

[56] N. Meshkov, A.J. Glick, and H.J. Lipkin. “Validity of many-body approximation methods for a solvable model: (II). Linearization procedures”.
In: *Nucl. Phys.* 62.2 (1965), pp. 199–210. DOI: 10.1016/0029-5582(65)90863-1.

[57] S.P. Mikheyev and A.Yu. Smirnov. “Resonance Amplification of Oscillations in Matter and Spectroscopy of Solar Neutrinos”.
In: *Sov. J. Nucl. Phys.* 42 (1985), pp. 913–917.

[58] C. Monroe et al.
“Programmable quantum simulations of spin systems with trapped ions”.
In: *Rev. Mod. Phys.* 93 (2 Apr. 2021), p. 025001.
DOI: 10.1103/RevModPhys.93.025001.

[59] Witold Nazarewicz. *Computational Problem Lipkin Model*.

[60] Feng Pan, J.P. Draayer, and W.E. Ormand.
“A particle-number-conserving solution to the generalized pairing problem”.
In: *Physics Letters B* 422 (1998), pp. 1–5. DOI: 10.1016/S0370-2693(98)00034-3.

[61] James T. Pantaleone. “Neutrino oscillations at high densities”.
In: *Phys. Lett. B* 287 (1992), pp. 128–132. DOI: 10.1016/0370-2693(92)91887-F.

[62] Y. Pehlivan et al. “Invariants of collective neutrino oscillations”.
In: *Phys. Rev. D* 84 (6 2011), p. 065008. DOI: 10.1103/PhysRevD.84.065008.

[63] Alberto Peruzzo et al.
“A variational eigenvalue solver on a photonic quantum processor”.
In: *Nat. Commun.* 5 (1 2014). DOI: 10.1038/ncomms5213.

[64] Google AI Quantum et al.
“Hartree-Fock on a superconducting qubit quantum computer”.
In: *Science* 369.6507 (2020), pp. 1084–1089. DOI: 10.1126/science.abb9811.

[65] Giulio Racah. “Theory of Complex Spectra. III”.
in: *Phys. Rev.* 63 (9-10 1943), pp. 367–382. DOI: 10.1103/PhysRev.63.367.

[66] R.W. Richardson and N. Sherman.
“Exact eigenstates of the pairing-force Hamiltonian”.
In: *Nucl. Phys.* 52 (1964), pp. 221–238. DOI: 10.1016/0029-5582(64)90687-X.

[67] Alessandro Roggero.
“Dynamical phase transitions in models of collective neutrino oscillations”.
In: *Phys. Rev. D* 104 (12 Dec. 2021), p. 123023.
DOI: 10.1103/PhysRevD.104.123023.

[68] Alessandro Roggero.
“Entanglement and many-body effects in collective neutrino oscillations”.
In: *Phys. Rev. D* 104 (10 Nov. 2021), p. 103016.
DOI: 10.1103/PhysRevD.104.103016.

[69] Alessandro Roggero et al.
“Preparation of excited states for nuclear dynamics on a quantum computer”.
In: *Phys. Rev. C* 102 (6 2020), p. 064624. DOI: 10.1103/PhysRevC.102.064624.

[70] Alessandro Roggero et al. “Quantum computing for neutrino-nucleus scattering”.
In: *Phys. Rev. D* 101 (7 2020), p. 074038. DOI: 10.1103/PhysRevD.101.074038.

[71] Ermal Rrapaj. “Exact solution of multiangle quantum many-body collective neutrino-flavor oscillations”. In: *Phys. Rev. C* 101 (6 2020), p. 065805.
DOI: 10.1103/PhysRevC.101.065805.

[72] Raymond F. Sawyer. “"Classical" instabilities and "quantum" speed-up in the evolution of neutrino clouds”.
In: *arXiv: High Energy Physics - Phenomenology* (2004).

[73] E. Schrödinger.
“An Undulatory Theory of the Mechanics of Atoms and Molecules”.
In: *Phys. Rev.* 28 (6 1926), pp. 1049–1070. DOI: 10.1103/PhysRev.28.1049.

[74] Jacob T. Seeley, Martin J Richard, and Peter J. Love. “The Bravyi-Kitaev transformation for quantum computation of electronic structure”.
In: *J. Chem. Phys.* 137 (2012). DOI: 10.1063/1.4768229.

[75] Peter W. Shor. “Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer”.
In: *SIAM Journal on Computing* 26.5 (1997). DOI: 10.1137/S0097539795293172.

[76] G. Sigl and G. Raffelt.
“General kinetic description of relativistic mixed neutrinos”.
In: *Nucl. Phys. B* 406.1 (1993), pp. 423–451.
DOI: 10.1016/0550-3213(93)90175-O.

[77] J. C. Slater. “The Theory of Complex Spectra”.
In: *Phys. Rev.* 34 (10 1929), pp. 1293–1322. DOI: 10.1103/PhysRev.34.1293.

[78] Masuo Suzuki. “General theory of fractal path integrals with applications to many-body theories and statistical physics”.
In: *J. Math. Phys.* 32.2 (1991), pp. 400–407. DOI: 10.1063/1.529425.

[79] IBM Q team. $ibmq_washington$. 2021. URL: https://quantum-computing.ibm.com/services/resources?type=Eagle&system=ibm_washington.

[80] Alexander Tichai.
“Many-Body Perturbation Theory for Ab Initio Nuclear Structure”.
PhD thesis. Technical University of Darmstadt, 2017.

[81] Minh C. Tran et al. “Faster Digital Quantum Simulation by Symmetry Protection”.
In: *PRX Quantum* 2 (1 2021), p. 010323. DOI: 10.1103/PRXQuantum.2.010323.

[82] H.F. Trotter. “On the product of semi-groups of operators”.

In: *Proc. Amer. Math. Soc.* 10 (1959).
DOI: 10.1090/S0002-9939-1959-0108732-6.

[83] Farrokh Vatan and Colin Williams.
"Optimal quantum circuits for general two-qubit gates".
In: *Phys. Rev. A* 69 (3 2004), p. 032315. DOI: 10.1103/PhysRevA.69.032315.

[84] Joel J. Wallman and Joseph Emerson.
"Noise tailoring for scalable quantum computation via randomized compiling".
In: *Phys. Rev. A* 94 (5 2016), p. 052325. DOI: 10.1103/PhysRevA.94.052325.

[85] Daochen Wang, Oscar Higgott, and Stephen Brierley.
"Accelerated Variational Quantum Eigensolver".
In: *Phys. Rev. Lett.* 122 (14 2019), p. 140504.
DOI: 10.1103/PhysRevLett.122.140504.

[86] G. C. Wick. "The Evaluation of the Collision Matrix".
In: *Phys. Rev.* 80 (2 1950), pp. 268–272. DOI: 10.1103/PhysRev.80.268.

[87] L. Wolfenstein. "Neutrino Oscillations in Matter".
In: *Phys. Rev. D* 17 (1978), pp. 2369–2374. DOI: 10.1103/PhysRevD.17.2369.

[88] William K. Wootters.
"Entanglement of Formation of an Arbitrary State of Two Qubits".
In: *Phys. Rev. Lett.* 80 (10 1998), pp. 2245–2248.
DOI: 10.1103/PhysRevLett.80.2245.

APPENDIX A

SU(2) COMMUTATION RELATIONS

In this appendix we prove the mapping between the angular momentum SU(2) operators (J_+, J_-, and J_z) and the fermionic creation and annihilation operators (a and $a^\dagger$). The mappings (4.10) and (4.11) are shown to hold through the demonstration that the SU(2) commutation relations (4.14) and (4.15) hold under the substitution of fermionic operators for SU(2) operators from the mappings. Here we define $j_{m\pm} = a^\dagger_{m\pm} a_{m\mp}$. We have that

$$\begin{aligned}
[J_+, J_-] &= [\sum_m j_{m+}, \sum_n j_{n-}] \\
&= \sum_{mn} [j_{m+}, a^\dagger_{n-} a_{n+}] \\
&= \sum_{mn} \left([j_{m+}, a^\dagger_{n-}]\, a_{n+} + a^\dagger_{n-} [j_{m+}, a_{n+}] \right) \\
&= \sum_{mn} \delta_{mn} \left(a^\dagger_{m+} a_{n+} - a^\dagger_{n-} a_{m-} \right) \\
&= \sum_m \left(a^\dagger_{m+} a_{m+} - a^\dagger_{m-} a_{m-} \right) \\
&= 2J_z, \qquad \text{(A.1)}
\end{aligned}$$

and

$$\begin{aligned}
[J_z, J_\pm] &= [\frac{1}{2} \sum_{m\sigma} \sigma a^\dagger_{m\sigma} a_{m\sigma}, \sum_n j_{n\pm}] \\
&= \frac{1}{2} \sum_{mn\sigma} \sigma [a^\dagger_{m\sigma} a_{m\sigma}, j_{n\pm}] \\
&= \frac{1}{2} \sum_{mn\sigma} \sigma \left(a^\dagger_{m\sigma} [a_{m\sigma}, j_{n\pm}] + [a^\dagger_{m\sigma}, j_{n\pm}]\, a_{m\sigma} \right) \\
&= \frac{1}{2} \sum_{mn\sigma} \sigma \left(\delta_{\sigma\pm} a^\dagger_{m\sigma} a_{n\mp} - \delta_{-\sigma\pm} a^\dagger_{n\pm} a_{m\sigma} \right)
\end{aligned}$$

$$= \pm \sum_m a^\dagger_{m\pm} a_{m\mp}$$
$$= \pm J_\pm, \tag{A.2}$$

where we've used the commutations of the SU(2) ladder operator and the fermionic operators which are derived below

$$\begin{aligned}
\left[j_{p\sigma}, a^\dagger_{q\tau}\right] &= \left[a^\dagger_{p\sigma} a_{p-\sigma}, a^\dagger_{q\tau}\right] \\
&= a^\dagger_{p\sigma}\left[a_{p-\sigma}, a^\dagger_{q\tau}\right] + \left[a^\dagger_{p\sigma}, a^\dagger_{q\tau}\right] a_{p-\sigma} \\
&= a^\dagger_{p\sigma}\left(\left\{a_{p-\sigma}, a^\dagger_{q\tau}\right\} - 2a^\dagger_{q\tau} a_{p-\sigma}\right) \\
&+ \left(\left\{a^\dagger_{p\sigma}, a^\dagger_{q\tau}\right\} - 2a^\dagger_{q\tau} a^\dagger_{p\sigma}\right) a_{p-\sigma} \\
&= \delta_{pq}\delta_{-\sigma\tau} a^\dagger_{p\sigma} - 2\left\{a^\dagger_{p\sigma}, a^\dagger_{q\tau}\right\} a_{p-\sigma} \\
&= \delta_{pq}\delta_{-\sigma\tau} a^\dagger_{p\sigma}
\end{aligned} \tag{A.3}$$

$$\begin{aligned}
\left[j_{p\sigma}, a_{q\tau}\right] &= -\left[a_{q\tau}, k_{p\sigma}\right] = -\left[k^\dagger_{p\sigma}, a^\dagger_{q\tau}\right]^\dagger \\
&= -\left[k_{p-\sigma}, a^\dagger_{q\tau}\right]^\dagger = -(\delta_{pq}\delta_{\sigma\tau} a^\dagger_{p-\sigma})^\dagger \\
&= -\delta_{pq}\delta_{\sigma\tau} a_{p-\sigma}.
\end{aligned} \tag{A.4}$$

APPENDIX B

NORMALIZED HARTREE-FOCK ANSATZ

In this appendix we derive the normalization of the Hartree-Fock ansatz. The definition of a coherent state, applied to the group SU(2), gives that the SU(2) coherent state takes the form

$$|\tau\rangle = e^{\zeta J_+ - \bar{\zeta} J_-} |0\rangle , \tag{B.1}$$

where $\xi = (\theta/2)e^{-i\phi}$. We wish to write the coherent state in the form

$$|\tau\rangle' = N e^{\tau J_+} |0\rangle , \tag{B.2}$$

which we will do through the use of the SU(2) generating function

$$\langle\tau| e^{\tau} e^{\alpha_0 J_0} e^{\alpha_- J_-} |\tau\rangle \tag{B.3}$$

$$= (1+|\tau|^2)^{-2J} \left[e^{\alpha_0/2}|\tau|^2 + e^{-\alpha_0/2}(\alpha_+\tau^* + 1)(\alpha_-\tau + 1) \right]^{2J} , \tag{B.4}$$

and the BCH equation

$$e^{\zeta J_+ - \bar{\zeta} J_-} = e^{\tau J_+} e^{\ln(1+|\tau|^2) J_0} e^{-\bar{\tau} J_-} . \tag{B.5}$$

If the two forms of the coherent state are equal, then

$$1 \tag{B.6}$$

$$= \langle\tau|\tau\rangle' \tag{B.7}$$

$$= N \langle\tau| e^{\tau J_+} |0\rangle \tag{B.8}$$

$$= N \langle\tau| e^{\tau J_+} e^{-(\zeta J_+ - \bar{\zeta} J_-)} |\tau\rangle \tag{B.9}$$

$$= N \langle\tau| e^{\ln(1+|\tau|^2) J_0} e^{\bar{\tau} J_-} |\tau\rangle \tag{B.10}$$

$$= (1+|\tau|^2)^{-2J}\left[e^{\ln(1+|\tau|^2)/2}|\tau|^2 + e^{-\ln(1+|\tau|^2)/2}(1+|\tau|^2)\right]^{2J} \tag{B.11}$$

$$= (1+|\tau|^2)^{-2J}\left[(1+|\tau|^2)^{\frac{1}{2}}|\tau|^2 + (1+|\tau|^2)^{-\frac{1}{2}}(1+|\tau|^2)\right]^{2J} \tag{B.12}$$

$$= (1+|\tau|^2)^{J}, \tag{B.13}$$

which implies that $N = (1+|\tau|^2)^{-J}$.

APPENDIX C

LIPKIN HAMILTONIAN FOR HARTREE-FOCK

In this appendix, we derive the matrix element $\langle\tau|H|\sigma\rangle$ where H is the Lipkin Hamiltonian and

$$|\tau\rangle = (1+|\tau|^2)^{-\frac{\Omega}{2}} e^{\tau J_+} |0\rangle\,, \tag{C.1}$$

$$|\sigma\rangle = (1+|\sigma|^2)^{-\frac{\Omega}{2}} e^{\sigma J_+} |0\rangle\,, \tag{C.2}$$

are two arbitrary $SU(2)$ coherent states. Here $|0\rangle = |J-J\rangle$, $\tau = \tan(\theta/2)e^{-i\phi}$, $\sigma = \tan(\theta'/2)e^{-i\phi'}$, and $\Omega = J/2$. Expanding

$$\langle\tau|\,H\,|\sigma\rangle = (1+|\sigma|)^{-\frac{\Omega}{2}}\,\langle\tau|\,He^{\sigma J_+}\,|0\rangle \tag{C.3}$$

$$= (1+|\tau|)^{\frac{\Omega}{2}}(1+|\sigma|)^{-\frac{\Omega}{2}}\,\langle\tau|\,He^{(\sigma-\tau)J_+}\,|\tau\rangle \tag{C.4}$$

$$= (1+|\tau|)^{\frac{\Omega}{2}}(1+|\sigma|)^{-\frac{\Omega}{2}} \tag{C.5}$$

$$\times\left[\epsilon\,\langle\tau|\,J_0e^{(\sigma-\tau)J_+}\,|\tau\rangle - \frac{1}{2}V\,\langle\tau|\,(J_+^2+J_-^2)e^{(\sigma-\tau)J_+}\,|\tau\rangle\right]. \tag{C.6}$$

Using the SU(2) generating function

$$\langle\tau|\,e^{\alpha_-J_-}e^{\alpha_0J_0}e^{\alpha_+J_+}\,|\tau\rangle \tag{C.7}$$

$$= (1+|\tau|^2)^{-\Omega}\left[e^{-\alpha_0/2} + e^{\alpha_0/2}(\bar\tau+\alpha_-)(\tau+\alpha_+)\right]^{\Omega}, \tag{C.8}$$

we calculate each term separately, starting with

$$\begin{aligned}
&\langle\tau|\,J_0e^{(\sigma-\tau)J_+}\,|\sigma\rangle \\
&= \frac{\partial}{\partial\alpha_0}\,\langle\tau|\,e^{\alpha_0J_0}e^{(\sigma-\tau)J_+}\,|\tau\rangle\Big|_{\alpha_0=0} \\
&= \frac{1}{(1+|\tau|^2)^{\Omega}}\frac{\partial}{\partial\alpha_0}\left[e^{-\alpha_0/2} + \bar\tau\sigma e^{\alpha_0/2}\right]^{\Omega}\Big|_{\alpha_0=0}
\end{aligned}$$

$$= \frac{\Omega/2}{(1+|\tau|^2)^\Omega}(\sigma\bar{\tau}e^{\alpha_0/2} - e^{-\alpha_0/2})(\sigma\bar{\tau}e^{\alpha_0/2} + e^{-\alpha_0/2})^{\Omega-1}\Big|_{\alpha_0=0}$$
$$= \frac{\Omega/2}{(1+|\tau|^2)^\Omega}(\sigma\bar{\tau} - 1)(\sigma\bar{\tau} + 1)^{\Omega-1}, \tag{C.9}$$

followed by

$$\begin{aligned}
&\langle\tau| J_+^2 e^{(\sigma-\tau)J_+} |\sigma\rangle \\
&= \frac{\partial^2}{\partial\alpha_+^2} \langle\tau| e^{(\alpha_+ + \sigma - \tau)J_+} |\tau\rangle \Big|_{\alpha_+=0} \\
&= \frac{1}{(1+|\tau|^2)^\Omega} \frac{\partial^2}{\partial\alpha_+^2} [1 + \bar{\tau}(\alpha_+ + \sigma)]^\Omega \Big|_{\alpha_+=0} \\
&= \frac{\Omega(\Omega-1)}{(1+|\tau|^2)^\Omega}(\bar{\tau}^2)[1 + \bar{\tau}(\alpha_+ + \sigma)]^{\Omega-2}\Big|_{\alpha_+=0} \\
&= \frac{\Omega(\Omega-1)}{(1+|\tau|^2)^\Omega}(\bar{\tau}^2)(1 + \bar{\tau}\sigma)^{\Omega-2},
\end{aligned} \tag{C.10}$$

followed by

$$\begin{aligned}
&\langle\tau| J_-^2 e^{(\sigma-\tau)J_+} |\sigma\rangle \\
&= \frac{\partial^2}{\partial\alpha_-^2} \langle\tau| e^{\alpha_- J_-} e^{(\sigma-\tau)J_+} |\tau\rangle \Big|_{\alpha_-=0} \\
&= \frac{1}{(1+|\tau|^2)^\Omega} \frac{\partial^2}{\partial\alpha_-^2} [1 + (\bar{\tau} + \alpha_-)\sigma]^\Omega \Big|_{\alpha_-=0} \\
&= \frac{\Omega(\Omega-1)}{(1+|\tau|^2)^\Omega}(\sigma^2)[1 + (\bar{\tau} + \alpha_-)\sigma]^{\Omega-2}\Big|_{\alpha_-=0} \\
&= \frac{\Omega(\Omega-1)}{(1+|\tau|^2)^\Omega}(\sigma^2)(1 + \bar{\tau}\sigma)^{\Omega-2}.
\end{aligned} \tag{C.11}$$

Combining yields

$$\begin{aligned}
&\langle\tau| H |\sigma\rangle = (\Omega/2)[(1+|\tau|^2)(1+|\sigma|^2)]^{-\frac{\Omega}{2}} \\
&\times \left[\epsilon(\sigma\bar{\tau} - 1)(\sigma\bar{\tau} + 1)^{\Omega-1} - V(\Omega-1)\left(\bar{\tau}^2 + \sigma^2\right)(1 + \bar{\tau}\sigma)^{\Omega-2}\right].
\end{aligned} \tag{C.12}$$

APPENDIX D

DICKE STATE LEMMAS

In this appendix, we inductively prove the following two lemmas involving Dicke states

$$\sum_{p=1}^{n} A_p^\dagger \left|D_k^n\right\rangle = \sqrt{(n-k)(k+1)}\left|D_{k+1}^n\right\rangle, \tag{D.1}$$

$$\sum_{p=1}^{n} A \left|D_k^n\right\rangle = \sqrt{(n-k+1)(k)}\left|D_{k-1}^n\right\rangle, \tag{D.2}$$

The base case for the first lemma (D.1) is given by the case $(n,k) = (1,0)$ below

$$\begin{aligned} A_1^\dagger \left|D_0^1\right\rangle &= A_1^\dagger \left|0_1\right\rangle \\ &= \left|1_1\right\rangle \\ &= \left|D_1^1\right\rangle \\ &= \sqrt{(1-0)(0+1)}\left|D_1^1\right\rangle. \end{aligned} \tag{D.3}$$

Similarly, the base case for the second lemma (D.2) is given by the case $(n,k) = (1,1)$ below

$$\begin{aligned} A_1 \left|D_1^1\right\rangle &= A_1 \left|1_1\right\rangle \\ &= \left|0_1\right\rangle \\ &= \left|D_0^1\right\rangle \\ &= \sqrt{(1-1+1)(1)}\left|D_1^1\right\rangle. \end{aligned} \tag{D.4}$$

We prove the first lemma first by assuming, as our induction hypothesis, that (D.1) is true and show that it still holds when we take $n \longrightarrow n+1$, as follows

$$\sum_{p=1}^{n+1} A_p^\dagger \left|D_k^{n+1}\right\rangle = \left(\sum_{p=1}^{n} A_p^\dagger + A_{n+1}^\dagger\right)\left(\sqrt{\frac{k}{n+1}}\left|D_{k-1}^n\right\rangle\left|1\right\rangle + \sqrt{\frac{n-k+1}{n+1}}\left|D_k^n\right\rangle\left|0\right\rangle\right), \tag{D.5}$$

where we took out the $n+1$ term from the sum and used the recursive definition of the Dicke state (5.117). Using the first base case (D.3) yields

$$\left(\sum_{p=1}^{n} A_p^\dagger\right)\left(\sqrt{\frac{k}{n+1}}\left|D_{k-1}^n\right\rangle|1\rangle+\sqrt{\frac{n-k+1}{n+1}}\left|D_k^n\right\rangle|0\rangle\right)+\sqrt{\frac{n-k+1}{n+1}}\left|D_k^n\right\rangle|1\rangle\,. \tag{D.6}$$

Applying the induction hypothesis (D.1) and then combining the first and last terms yields

$$(k+1)\sqrt{\frac{n-k+1}{n+1}}\left|D_{k-1}^n\right\rangle|1\rangle+\sqrt{(n-k)(k+1)}\sqrt{\frac{n-k+1}{n+1}}\left|D_k^n\right\rangle|0\rangle \tag{D.7}$$

$$=\sqrt{(n-k+1)(k+1)}\left(\sqrt{\frac{k+1}{n+1}}\left|D_{k-1}^n\right\rangle|1\rangle+\sqrt{\frac{n-k}{n+1}}\left|D_k^n\right\rangle|0\rangle\right) \tag{D.8}$$

$$=\sqrt{(n-k+1)(k+1)}\left|D_{k+1}^{n+1}\right\rangle, \tag{D.9}$$

which completes the proof.

We now prove the second lemma first by assuming, as our induction hypothesis, that (D.2) is true and show that it still holds when we take $n \to n+1$, as follows

$$\sum_{p=1}^{n+1} A_p\left|D_k^{n+1}\right\rangle=\left(\sum_{p=1}^{n} A_p+A_{n+1}\right)\left(\sqrt{\frac{k}{n+1}}\left|D_{k-1}^n\right\rangle|1\rangle+\sqrt{\frac{n-k+1}{n+1}}\left|D_k^n\right\rangle|0\rangle\right), \tag{D.10}$$

where we took out the $n+1$ term from the sum and used the recursive definition of the Dicke state (5.117). Using the second base case (D.4) yields

$$\left(\sum_{p=1}^{n} A_p\right)\left(\sqrt{\frac{k}{n+1}}\left|D_{k-1}^n\right\rangle|1\rangle+\sqrt{\frac{n-k+1}{n+1}}\left|D_k^n\right\rangle|0\rangle\right)+\sqrt{\frac{k}{n+1}}\left|D_{k-1}^n\right\rangle|0\rangle \tag{D.11}$$

Applying the induction hypothesis (D.2) and then combining the first and last terms yields

$$\sqrt{k(n-k+2)}\sqrt{\frac{k-1}{n+1}}\left|D_{k-2}^n\right\rangle|1\rangle+(n-k+2)\sqrt{\frac{k}{n+1}}\left|D_{k-1}^n\right\rangle|0\rangle \tag{D.12}$$

$$=\sqrt{k(n-k+2)}\left(\sqrt{\frac{k-1}{n+1}}\left|D_{k-2}^n\right\rangle|1\rangle+\sqrt{\frac{n-k+2}{n+1}}\left|D_{k-1}^n\right\rangle|0\rangle\right) \tag{D.13}$$

$$=\sqrt{k(n-k+2)}\left|D_{k-1}^{n+1}\right\rangle, \tag{D.14}$$

which completes the proof.

APPENDIX E

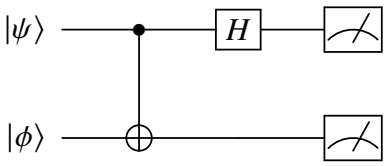
STATE-OVERLAP ALGORITHM

In this appendix we explain the efficient state-overlap algorithm that is used to find the eigenenergies of excited states. Consider two states $|\psi\rangle = a\,|0\rangle + b\,|1\rangle$ and $|\phi\rangle = c\,|0\rangle + d\,|1\rangle$. The overlap between these two states is

$$|\langle\psi|\phi\rangle|^2 = (a^*c + b^*d)\,(ac^* + bd^*) \tag{E.1}$$

$$= |a|^2|c|^2 + |b|^2|d|^2 + a^*bcd^* + ab^*c^*d. \tag{E.2}$$

Now consider the circuit



(E.3)

The progression of the state through the above circuit is

$$|\psi\phi\rangle = ac\,|00\rangle + ad\,|01\rangle + bc\,|10\rangle + bd\,|11\rangle \tag{E.4}$$

$$\rightarrow ac\,|00\rangle + ad\,|01\rangle + bc\,|11\rangle + bd\,|10\rangle \tag{E.5}$$

$$\rightarrow \frac{1}{\sqrt{2}}\big[(ac + bd)\,|00\rangle + (ad + bc)\,|01\rangle \tag{E.6}$$

$$+ (ac - bd)\,|10\rangle + (ad - bc)\,|11\rangle\,\big]. \tag{E.7}$$

The probability that $|\psi\phi\rangle = |11\rangle$ is

$$P_{11} = \frac{1}{2}(ad - bc)(ad - bc)^* \tag{E.8}$$

$$= \frac{1}{2}\left[|a|^2|d|^2 + |b|^2|c|^2 - ab^*c^*d - a^*bcd^*\right]. \tag{E.9}$$

Note that

$$P_{00} + P_{01} + P_{10} - P_{11} \tag{E.10}$$

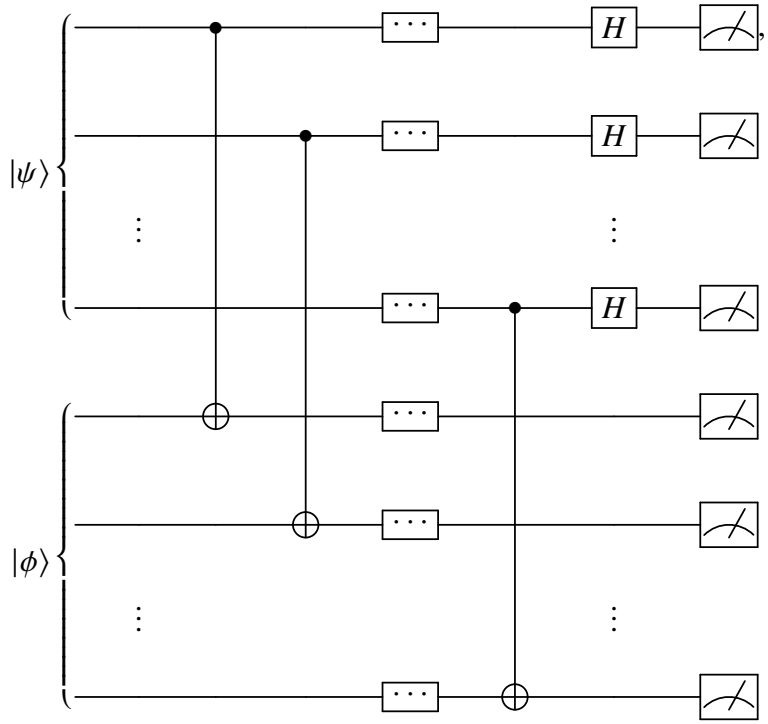
$$= 1 - 2P_{11} \tag{E.11}$$

$$= 1 - \left[|a|^2|d|^2 + |b|^2|c|^2 - ab^*c^*d - a^*bcd^*\right] \tag{E.12}$$

$$= |a|^2|c|^2 + |b|^2|d|^2 + ab^*c^*d + a^*bcd^* \tag{E.13}$$

$$= |\langle\psi|\phi\rangle|^2, \tag{E.14}$$

where we've used the fact that $1 = |a|^2 + |b|^2$. This can be generalized to



(E.15)

as explained in [36].

APPENDIX F

PAIR COMMUTATION RELATIONS

In this appendix, we use the fermionic anti-commutation relations (2.39) to prove the pair fermionic commutation relations (5.6-5.8). The first commutation relation is proved as follows

$$[A_p, A_q^\dagger] = [a_{p-}a_{p+}, a_{q+}^\dagger a_{q-}^\dagger] \tag{F.1}$$

$$= a_{p-}a_{q+}^\dagger [a_{p+}, a_{q-}^\dagger] + a_{p-}[a_{p+}, a_{q+}^\dagger]\, a_{q-}^\dagger$$
$$+ a_{q+}^\dagger [a_{p-}, a_{q-}^\dagger]\, a_{p+} + [a_{p-}, a_{q+}^\dagger]\, a_{q-}^\dagger a_{p+} \tag{F.2}$$

$$= a_{p-}a_{q+}^\dagger \cancel{\{a_{p+}, a_{q-}^\dagger\}} - 2a_{p-}a_{q+}^\dagger a_{q-}^\dagger a_{p+}$$
$$+ a_{p-}\{a_{p+}, a_{q+}^\dagger\}\, a_{q-}^\dagger - 2a_{p-}a_{q+}^\dagger a_{p+} a_{q-}^\dagger$$
$$+ a_{q+}^\dagger \{a_{p-}, a_{q-}^\dagger\}\, a_{p+} - 2a_{q+}^\dagger a_{q-}^\dagger a_{p-} a_{p+}$$
$$+ \cancel{\{a_{p-}, a_{q+}^\dagger\}}\, a_{q-}^\dagger a_{p+} - 2a_{q+}^\dagger a_{p-} a_{q-}^\dagger a_{p+} \tag{F.3}$$

$$= \delta_{pq}(a_{p-}a_{q-}^\dagger + a_{q+}^\dagger a_{p+})$$
$$- 2a_{p-}a_{q+}^\dagger \cancel{\{a_{q-}^\dagger, a_{p+}\}} - 2a_{q+}^\dagger \{a_{q-}^\dagger, a_{p-}\}\, a_{p+} \tag{F.4}$$

$$= \delta_{pq}(a_{p-}a_{q-}^\dagger - a_{q+}^\dagger a_{p+}) \tag{F.5}$$

$$= \begin{cases} a_{p-}a_{p-}^\dagger - a_{p+}^\dagger a_{p+} & \text{if } p = q \\ 0 & \text{if } p \neq q \end{cases}, \tag{F.6}$$

which we now compare to

$$\delta_{pq}(1 - N_p) = \delta_{pq}(\delta_{p-p-} - a_{p+}^\dagger a_{p+} - a_{p-}^\dagger a_{p-}) \tag{F.7}$$

$$= \delta_{pq}(a_{p-}^\dagger a_{p-} + a_{p-}a_{p-}^\dagger - a_{p+}^\dagger a_{p+} - a_{p-}^\dagger a_{p-}) \tag{F.8}$$

$$= \delta_{pq}(a_{p-}a_{p-}^\dagger - a_{p+}^\dagger a_{p+}) \tag{F.9}$$

$$= \begin{cases} a_{p-}a^\dagger_{p-} - a^\dagger_{p+}a_{p+} & \text{if } p = q \\ 0 & \text{if } p \neq q \end{cases}, \tag{F.10}$$

and note that they match, thus proving the first pair anti-commutation relation (5.6). The second pair anti-commutation relation is proved as follows

$$[N_p, A^\dagger_q] = [\sum_\sigma a^\dagger_{p\sigma}a_{p\sigma}, a^\dagger_{q+}a^\dagger_{q-}] \tag{F.11}$$

$$= \sum_\sigma [a^\dagger_{p\sigma}a_{p\sigma}, a^\dagger_{q+}a^\dagger_{q-}] \tag{F.12}$$

$$= \sum_\sigma a^\dagger_{p\sigma}a^\dagger_{q+}[a_{p\sigma}, a^\dagger_{q-}] + a^\dagger_{p\sigma}[a_{p\sigma}, a^\dagger_{q+}]\, a^\dagger_{q-} + a^\dagger_{q+}[a^\dagger_{p\sigma}, a^\dagger_{q-}]\, a_{p\sigma} + [a^\dagger_{p\sigma}, a^\dagger_{q+}]\, a^\dagger_{q-}a_{p\sigma} \tag{F.13}$$

$$= \sum_\sigma (a^\dagger_{p\sigma}a^\dagger_{q+}\{a_{p\sigma}, a^\dagger_{q-}\} - 2a^\dagger_{p\sigma}a^\dagger_{q+}a^\dagger_{q-}a_{p\sigma} + a^\dagger_{p\sigma}\{a_{p\sigma}, a^\dagger_{q+}\}\, a^\dagger_{q-} - 2a^\dagger_{p\sigma}a^\dagger_{q+}a_{p\sigma}a^\dagger_{q-} + a^\dagger_{q+}\{a^\dagger_{p\sigma}, a^\dagger_{q-}\}\, a_{p\sigma} - 2a^\dagger_{q+}a^\dagger_{q-}a^\dagger_{p\sigma}a_{p\sigma} + \{a^\dagger_{p\sigma}, a^\dagger_{q+}\}\, a^\dagger_{q-}a_{p\sigma} - 2a^\dagger_{q+}a^\dagger_{p\sigma}a^\dagger_{q-}a_{p\sigma}) \tag{F.14}$$

$$= \sum_\sigma [\delta_{pq}(\delta_{\sigma+}a^\dagger_{p\sigma}a^\dagger_{q-} + \delta_{\sigma-}a^\dagger_{p\sigma}a^\dagger_{q+}) - 2(a^\dagger_{p\sigma}a^\dagger_{q+}\{a^\dagger_{q-}, a_{p\sigma}\} + a^\dagger_{q+}\{a^\dagger_{q-}, a^\dagger_{p\sigma}\}\, a_{p\sigma})] \tag{F.15}$$

$$= \delta_{pq} \sum_\sigma [\delta_{\sigma+}a^\dagger_{p\sigma}a^\dagger_{q-} + \delta_{\sigma-}a^\dagger_{p\sigma}a^\dagger_{q+} - 2\delta_{\sigma-}a^\dagger_{p\sigma}a^\dagger_{q+}] \tag{F.16}$$

$$= \delta_{pq}(a^\dagger_{p+}a^\dagger_{q-} + a^\dagger_{p-}a^\dagger_{q+} - 2a^\dagger_{p-}a^\dagger_{q+}) \tag{F.17}$$

$$= \delta_{pq}(a^\dagger_{p+}a^\dagger_{q-} - a^\dagger_{p-}a^\dagger_{q+}) \tag{F.18}$$

$$= \begin{cases} 2a^\dagger_{p+}a^\dagger_{p-} & \text{if } p = q \\ 0 & \text{if } p \neq q \end{cases}, \tag{F.19}$$

since $\{a^\dagger_{p+}, a^\dagger_{p-}\} = 0$, which we now compare to

$$2\delta_{pq}A^\dagger_p = 2\delta_{pq}a^\dagger_{p+}a^\dagger_{p-} \tag{F.20}$$

$$= \begin{cases} 2a^\dagger_{p+}a^\dagger_{p-} & \text{if } p = q \\ 0 & \text{if } p \neq q \end{cases}, \tag{F.21}$$

and note that they match, thus proving the second pair anti-commutation relation (5.7). The third anti-commutation relation (5.8) can be derived from the second (5.7) as follows

$$\left[N_p, A_q\right] = -[N_p, A^\dagger_q]^\dagger \tag{F.22}$$

$$= -2\delta_{pq}A^\dagger_p, \tag{F.23}$$

since $N^\dagger_p = N_p$.