

Master's Thesis (Academic Year 2025)

Breaking symmetry for faster simulation of strongly correlated systems

Keio University
Graduate School of Media and Governance

Pawan Sharma Poudel

January 2026

Abstract of Master's Thesis of Academic Year 2025

Breaking symmetry for faster simulation of strongly correlated systems

In this work we use a widely used concept in classical chemistry calculations, symmetry-breaking, to speed-up the ground state energy estimation task in molecular systems using the Quantum Imaginary Time Evolution (QITE) algorithm. The Hartree—Fock (HF) wave function, commonly used for approximating molecular ground states, becomes nonideal in strongly correlated molecules due to the inherent multiconfigurational nature of the wave function, limiting accuracy in QITE. The speed of convergence towards the ground state in QITE is inversely proportional to the energy gap between the ground state and the first excited state. We introduce spin and spatial symmetry broken wave function and raise the energy gap between the ground state and the first excited state to enable faster convergence towards the lowest energy eigenstate in QITE. We verify the approach with numerical emulations, and gate-based simulations of H₂ molecule (tested also on IBM real machine), and a square tetrahydrogen cluster using measurement-assisted unitary approximation. We further optimize the state tomography in QITE, improving the sampler based simulation speed.

QITE demonstrates faster convergence to the ground state with broken symmetry (BS) compared to HF, particularly after the molecule exhibits a certain diradical character specific to the molecule. Prior to this point, HF remains more effective, suggesting a transition threshold of diradical character for wave function selection. Additionally, the overlap analysis with Complete Active Space Configuration Interaction (CAS-CI) wave function shows that BS has a larger initial overlap than HF in higher-spin, multi-configurational systems like triple bond dissociation in N₂ molecule. This method provides a pathway for improved energy simulations in open shell systems, where wave function accuracy significantly impacts downstream quantum algorithms and practical applications in quantum chemistry.

Keywords

Broken symmetry wave function, Quantum Imaginary Time Evolution, Spin operator, Ground state energy estimation

Graduate School of Media and Governance
Keio University

Pawan Sharma Poudel

Acknowledgment

I am very grateful towards my supervisor Professor Rodney Van Meter and feel extremely fortunate to be his student. My journey in research formally started under him and his guidance has always been the map I follow in research to give back to the society. I put my other supervisors, Michal Hajdušek and Kenji Sugisaki on the top who have guided me through this journey in quantum computational chemistry.

Rod has always been my guardian at Keio, taking care of every necessities I would ask him for during the course of research and beyond. I could not ask for a better advisor than him, and he will always remain my role model. Michal is the person you should know about if you are reading this thesis. His art of explaining every concepts of quantum mechanics to quantum information is something I have been a fan of. I got an opportunity to spend some research time with him and learn a lot during that span. Sugisaki-sensei has been an integral part of my research journey, mentoring me while navigating through complex chemistry concepts. This work would not have been possible without his guidance and patience answering my dumb questions.

I would like to thank my thesis advisory committee, Rodney Van Meter, Michal Hajdušek and Hideyuki Kawashima. Their comments and suggestions have been integral in preparing this thesis. In addition, suggestions from Jin Mitsugi and other RG faculty were largely valuable.

I am very thankful towards Naphan Benchasattabuse at AQUA, who continuously guided me through navigating quantum computing, being a very good senior. I have been very fortunate to share beautiful memories and grow together with Hikaru Yokomori, Natchapol Patamawisut, Gordon Cui, and all good hearts in AQUA lab. I would want to thank these people for joining to try every restaurants in Shonandai, watch late night football, organize Yakiniku at Hikaru's place, travel together to conferences and make life full of fun during the Masters.

I want to thank Naoki Yamamoto and Keio Quantum Computing Center (KQCC) for giving me opportunity to present and discuss quantum chemistry research.

I would like to acknowledge Quantum-LEAP Flagship Program and JST Moonshot R&D for funding this research. I express my gratitude towards NEC Capital Solutions (now CRTS) for arranging flexible working hours and allowing me to study continuing hybrid work at the workplace.

I should not miss Professor Ved Kafle who guided me to apply for Keio University and let me know more about Rod.

Of all, there lie my nearest friends, who have always been my side throughout my journey from bachelors to now. Thank you Basanta Rijal, Aasman Bashyal, Sushil Raj Regmi, Pranish Acharya, Nischal Budhathoki, Kamal Godar, Ujjwal Pokharel, Aashish Adhikari, Praveen Gautam, Prashant Ghimire, Sandip Adhikari, Rishav Dhungel dai, Nischal Nepal dai, Kishor Sapkota, Ankit Kharel, Sandeep Poudel, Pradipta Pradhanang, Tara Dhakal didi and everyone I talked, interacted and shared my time with.

Ending, I remain forever indebted to my parents, Kamala Dhakal and Purnananda Sharma Poudel, my brother Prabin Sharma Poudel, and my girlfriend Sonia Tiwari. Their unwavering love, patience, and belief in me have been the quiet strength behind every step of my journey. Everything I am and everything I have achieved is rooted in their care and sacrifice, and I consider myself profoundly blessed to share my life with them.

Contents

Acknowledgment	ii
1 Introduction	1
1.1 Research contribution	2
1.2 Thesis structure	2
2 Theory	4
2.1 Qubits	4
2.1.1 Quantum state	4
2.1.2 Bloch sphere	6
2.1.3 Multiple qubits	6
2.2 Quantum properties	7
2.2.1 Superposition	7
2.2.2 Entanglement	8
2.3 Quantum computation	10
2.3.1 Measurement	10
2.3.2 Density operator	11
2.3.3 Quantum gates	11
2.4 A quick visit to molecular quantum mechanics	14
2.4.1 Wave function	14
2.4.2 Schrödinger equation	15
2.4.3 Spin operators	16
2.4.4 Atomic and molecular orbital theory	19
2.4.5 Electronic structure problem	24
2.4.6 Hartree–Fock theory	31
2.4.7 Density functional theory	39
2.4.8 Strongly correlated systems	40
2.5 Quantum simulation	42
2.5.1 Suzuki–Trotter simulation	43
2.5.2 qDrift	44
2.5.3 Other methods	47
2.6 Quantum imaginary time evolution	47
2.6.1 Imaginary time evolution	47
2.6.2 Unitary approximation	47
3 Design rationale and implementation	50
3.1 Motivation	50

3.1.1	Problems with simulating strongly-correlated systems	50
3.2	The broken-symmetry wave function	51
3.3	Spin penalty to Hamiltonian	53
3.3.1	Intuition behind symmetry breaking and penalty	54
3.4	Algorithmic steps	55
3.4.1	Parameters initialization	57
3.4.2	Hamiltonian construction	57
3.4.3	Initial state preparation	57
3.4.4	QITE loop	58
3.5	Computational conditions	63
4	Results and discussion	65
4.1	ITE with matrix exponentiation	65
4.2	Simulator results	67
4.2.1	Numerical simulation with statevector	67
4.2.2	Simulation with sampler	70
4.3	Optimizing tomography for simulating larger systems	71
4.4	Results on IBM quantum device	74
4.5	Initial fidelity in strongly correlated region	76
4.6	QITE algorithmic complexity	77
5	Conclusion and future direction	81
5.1	Conclusion	81
5.2	Future works	82
5.3	Code Availability	82

List of Figures

2.1	Norms of a quantum state	6
2.2	Bloch sphere	7
2.3	Single-qubit gates and their corresponding matrix representations	12
2.4	Hadamard gate and its matrix representation	12
2.5	Controlled-NOT (CNOT) gate	13
2.6	Toffoli gate	14
2.7	Spin cone diagram	18
2.8	Orbital filling diagram	21
2.9	Carbon atom electronic configuration	22
2.10	Molecular orbitals from s orbitals	23
2.11	Molecular orbitals from p_z orbitals	23
2.12	Molecular orbitals from p_x orbitals	24
2.13	Molecular orbitals from p_y orbitals	25
2.14	Molecular orbitals in N_2 molecule	26
2.15	Comparison of two orbital occupation configurations across the three p -orbitals	37
2.16	H_2 molecule singlet ground state configuration showing antiparallel spins in the bonding σ_u orbital	37
2.17	Open-shell singlet state as a superposition of spin configurations	38
2.18	Low-spin triplet state as a superposition of spin configurations	38
2.19	Superposition of singlet and triplet states	38
2.20	Circuit for simulating the Hamiltonian given as $\alpha X \otimes X$	43
3.1	Branching diagram for the construction of spin eigenfunctions	52
3.2	Active orbitals of the N_2 molecule	53
3.3	Energy gap comparison of two lowest energy states for HF and BS with penalty	55
3.4	Flowchart of Quantum Imaginary Time Evolution algorithm	56
3.5	Electron spin diagrams and corresponding quantum circuit implementations for RHF and BS states of P4	58
3.6	Subroutine for a single QITE step	60
3.7	Subroutine for measuring energy	60
3.8	Subroutine for calculating the expectation value	61
3.9	Subroutine for applying Pauli rotations	61
3.10	Subroutine and its circuit realizations for applying Pauli rotation to qubits .	62
4.1	ITE by direct matrix exponentiation and RHF-BS comparison on a N_2 molecule	67

4.2	QITE with numerical simulation and RHF-BS comparison of H ₂ molecule	68
4.3	QITE with numerical simulation and RHF-BS comparison of P4 cluster system	69
4.4	Sampler based QITE implementation and RHF-BS comparison for H ₂ molecule	70
4.5	Optimized QITE algorithm with grouped Hamiltonian terms	72
4.6	Subroutine for processing one group of commuting Hamiltonian terms efficiently	73
4.7	Sampler based optimized QITE simulation of P4 cluster	74
4.8	QITE implementation and RHF-BS comparison of H ₂ molecule on a quantum processor	75
4.9	IBM Heron r2 processor qubit arrangement	76
4.10	Initial fidelities of RHF-BS wave functions of N ₂ molecule	77

List of Tables

2.1	Spin states and their quantum numbers for two-spin systems	19
4.1	Complexity table for computationally intensive blocks	79
4.2	Gate-type decomposition with complexity analysis	79

Chapter 1

Introduction

Simulating quantum many-body systems accurately is one of the central challenges in physics, chemistry and material science. The behavior of systems is governed by the Schrödinger equation, whose solution requires exponential increase in resources as the system size increases. As a result, the classical computational resources and techniques become untractable for larger and strongly correlated systems. This suggests the need of a new paradigm of computing. Quantum simulation provides a promise of exponential speed-ups to problems of simulating many-body systems which would take thousands or millions of years for the largest supercomputer in the world to compute [1]. As most of the physical systems of interest can be described by their Hamiltonian, we can simulate the quantum evolution of those system under their Hamiltonian.

A molecular system lives is the most stable in its lowest energy state called ground state. Finding the ground state of a system is among the various important problems in Hamiltonian simulation. This problem carries more attention due to its use in drug discovery [2], semiconductor and battery manufacturers [3], and researchers including chemists, physicists and computational scientists.

Various quantum algorithmic approaches like Quantum Phase Estimation [1], [4], [5], Variational Quantum Eigensolver (VQE) [6], Quantum Monte Carlo (QMC) [7], Adiabatic State Preparation (ASP) [1], and Quantum Imaginary Time Evolution (QITE) [8] are proposed to solve a common problem of finding the ground state energy. Despite lots of improvements to date, the application of these quantum algorithms to strongly correlated systems, such as multinuclear transition metal complexes and molecules under covalent bond dissociation, remains a challenge [9], [10], [11]. Under the strong correlation, the ground state wave function shows strong multi-configurational character, and the Hartree–Fock (HF) wave function is not a good approximation of the ground state. The small energy gap between the ground and excited states is another important feature of the strong correlation, and ASP and QITE usually suffer from slow convergence. Acceleration of the ground state energy estimation (GSEE) [12] for strongly correlated molecules

through improving the convergence behavior is an important issue and I use the concept of symmetry-breaking to tackle it in this thesis.

1.1 Research contribution

We address the problem of slow convergence of QITE for molecules by proposing a new variant that takes into account the open shell character of a molecule throughout its potential energy surface (PES) during bond dissociation. We have figured out that the introduction of spin penalty to the Hamiltonian during QITE, started with broken-symmetry (BS) wave function, provides an improvement to the time it takes to reach the ground state. We demonstrate our approach to matrix exponentiation for the time evolution of BS wave function with ITE on the case of N₂ molecule. As a proof of concept, to show that it performs well with unitary approximation as in QITE, we perform numerical emulations for the H₂ molecule PES and square tetrahydrogen (P4) cluster [13] with interatomic distances being 2.0 Bohr \approx 1.058 Å. We identify the region in PES of H₂ and N₂ molecules where our method offers a significant speedup in the convergence rate. The region is identified using the diradical character y as a measure of open shell electronic configuration, which we use to select the appropriate starting wave function for QITE.

I further carry out gate based simulations on a simulator in Qiskit [14] and on real IBM quantum device. While the tomographic step consumes a large number of measurement circuits and has higher shots requirements, I optimize the basis requirements for tomography and reduce the number of measurement circuits for simulation of larger molecular systems. Towards the end of the thesis, I have put some research questions that can be looked into to explore the regime of ground state energy estimation with QITE as the system size grows.

1.2 Thesis structure

This thesis is structured to give the readers an easy to understand outlook of the concepts of quantum information processing and molecular quantum mechanics.

Chapter 2 gives background on qubits, quantum properties, observables and measurement operations. Section 2.4 discusses the quantum mechanical picture of describing a state, and preliminary methods used in state preparation for running on a quantum computer. I provide a detailed literature on Hartree–Fock molecular orbital theory around which our research contribution stays. Then it goes on to describing methods used in quantum simulation which are implemented in this work. I expect the readers to have a basic understanding of valence-bond theory taught in high-school chemistry lectures.

In chapter 3, I first highlight the motivation behind this work, and present the problems in simulating strongly correlated system. Then I present the algorithmic flow and touch on how each of the steps can be implemented on a gate based quantum computer. In addition, I have shown in detail the construction of a symmetry broken wave function and how it

guarantees our claim. Further, the computational conditions and considered constraints for simulation are discussed in section 3.5.

Chapter 4 details our results on different molecules. The results include those with matrix exponentiation using ITE, numerical simulation and gate based simulation on simulator with QITE and simulation on a real quantum device. I also present the techniques to optimize the state tomography process and the quantum gate complexity in executing QITE on a sampler based simulator.

Chapter 5 concludes the work with very important insights on how this work can be extended for quantum speed-ups and tested for benchmarking molecules. I also open source the source code used in this project.

Chapter 2

Theory

In this chapter, I put light on the theoretical foundation that is important in understanding the overall aspect of this thesis.

2.1 Qubits

Coined for the first time by Schumacher in 1995 in his famous work on Quantum Coding [15], the term *qubit* refers to quantum-bit. In contrast to the binary bits we are familiar with the digital computers we use, qubits hold interesting properties that can describe quantum phenomena like superposition and entanglement. Having been used in the literature for a long time, in the entirety of this thesis we term the general existing computers as of the date of writing this thesis as classical computers and the basis of the computing based on transistors and binary bits as classical computing. While the classical bits can store and process information as 0 or 1, quantum bits allow us to process information as $|0\rangle$, $|1\rangle$ and everything between $|0\rangle$ and $|1\rangle$. This property forms the basis of advantage quantum computers are meant to offer over the classical computers.

2.1.1 Quantum state

A quantum state is described as, $|\Psi\rangle$. The notation “ $|\ \rangle$ ” is called *dirac ket-notation*, or just *ket*. The state $|\Psi\rangle$ is commonly pronounced as, “ket-psi.” Two very common quantum states, analogous to classical 1 and 0 are $|1\rangle$ and $|0\rangle$ respectively. In vector notation, $|0\rangle$ and $|1\rangle$ are written as $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$ respectively. Similarly, algebraically, the complex conjugate transpose of this ket notation is a bra, “ $\langle |$,” which is a dual vector. The dual vector corresponding to a vector $|u\rangle$ is $\langle u| = |u\rangle^\dagger = ((|u\rangle)^*)^T$.

An arbitrary quantum state $|\Psi\rangle$ can be represented as

$$|\Psi\rangle = \alpha|\phi_0\rangle + \beta|\phi_1\rangle, \quad (2.1)$$

where α and β are the probability amplitudes, and states $|\phi_0\rangle$ and $|\phi_1\rangle$ are the orthonormal basis vectors that form the Hilbert space, where the state $|\Psi\rangle$ lives. Since the coefficients α and β are probability amplitudes, their absolute squares sum to 1; $|\alpha|^2 + |\beta|^2 = 1$. This guarantees that the state $|\Psi\rangle$ is normalized.

A quantum state corresponding to n -qubits is a 2^n dimensional Hilbert space where these qubits can live. A *Hilbert space* is also known as inner-product space in a finite-dimensional complex vector space with inner products and norm. Hilbert space satisfies the completeness property that states that every Cauchy (some uniform convergent) sequence converges within the same complex inner product space. For example, we can take a counter example of a space which is not a Hilbert space. Consider a polynomial space $\mathbf{P}_\infty(x)$ with a Cauchy sequence: $x^0 + \frac{x^1}{1!} + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$. This sequence converges to e^x which does not lie in $\mathbf{P}_\infty(x)$ while $\{x^0, x^1, x^2, x^3, \dots\}$ lie on the polynomial space. Thus, the polynomial space is not a Hilbert space.

(1) Inner product

At this point, we should understand the inner product of two quantum states described by $|u\rangle$ and $|v\rangle$ as $\langle u|v\rangle$. Since it is product of a row and a column matrix, the inner product gives a complex scalar quantity. In quantum information, the inner product of vectors $|u\rangle$ and $|v\rangle$ can be denoted as, $(|u\rangle, |v\rangle)$ or $\langle u|v\rangle$ or $\langle u, v\rangle$ and obtained as,

$$\langle u|v\rangle = \text{Tr}(|u\rangle\langle v|) = \text{Tr}(|v\rangle\langle u|) \in \mathbb{C}. \quad (2.2)$$

Here, Tr refers to the trace, which mathematically is the sum of main-diagonal elements. While we take the inner product of two states, $|u\rangle$ and $|v\rangle$, as $\langle u|v\rangle$, we should understand this intuitively as the overlap amplitude of the state $|v\rangle$ along the direction of $|u\rangle$.

Thus, this takes us to a postulate of quantum mechanics that states,

“For a system prepared in state $|v\rangle$, the absolute square of the inner product of $|u\rangle$ and $|v\rangle$, i.e., $|\langle u|v\rangle|^2$ gives the probability of finding the state $|v\rangle$ in state $|u\rangle$. ”

The reader is suggested to refer to section 2.1.4 of *Quantum Computation and Quantum Information* [16] for the criteria that have to be satisfied for a function (\cdot, \cdot) mapping from an inner product space $V \times V$ to \mathbb{C} , to be an inner product.

(2) Norm

The norm of a state $|u\rangle$, which we have used above to refer the Euclidean norm or 2-norm is the square root of the inner product of the state with itself,

$$\| |u\rangle \| = \sqrt{\langle u|u\rangle} \in \mathbb{R} \geq 0. \quad (2.3)$$

All unitary operations acting on a quantum state should preserve the Euclidean norm of that state.

In quantum mechanics, we discuss other norms too, based on the scenario. One is 1-norm, which is given by

$$\| |u\rangle \|_1 = \sum_i |c_i|, \quad (2.4)$$

where c_i are the coefficients/amplitudes of the quantum state vector $|u\rangle$.

Similarly, infinity-norm (∞ -norm) is obtained as the maximum of the amplitudes, given as

$$\| |u\rangle \|_\infty = \max_i |c_i|. \quad (2.5)$$

Figure 2.1 shows the geometric representation of different norms. While the diagonal side is the 2-norm, referring to the distance between two points, the largest component after projection is the infinity-norm.

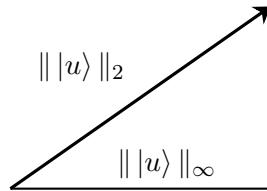


Figure 2.1: Geometric visualization of the ℓ_2 and ℓ_∞ norms of a quantum state vector $|u\rangle$.

2.1.2 Bloch sphere

Bloch sphere is a visual representation of the state of a single qubit. Since a quantum state lives in a finite dimensional Hilbert space and needs an orthonormal basis to describe it, the Bloch sphere provides a clear geometric picture of the state and its rotation about different basis of consideration.

Angles θ and ϕ determine the point on the surface of the Bloch sphere. As long as the state $|\Psi\rangle$ is normalized, it stays on the surface. If we do not consider the global phase (which we normally do, as it has no physical effect on any observable), any state represented in the Bloch sphere is represented as

$$|\Psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\phi}\sin\frac{\theta}{2}|1\rangle. \quad (2.6)$$

2.1.3 Multiple qubits

Following the pattern we had for single qubit representation with dirac notation, we represent multiple qubits as, $|00\rangle$, $|01\rangle$, $|10\rangle$, $|11\rangle$ in computational basis. Here, by computational basis, I mean that this basis set $\{00, 01, 10, 11\}$ is the most natural basis and is the most widely used in the literature of quantum mechanics. Yet the $\{++, +-, -+, --\}$ basis

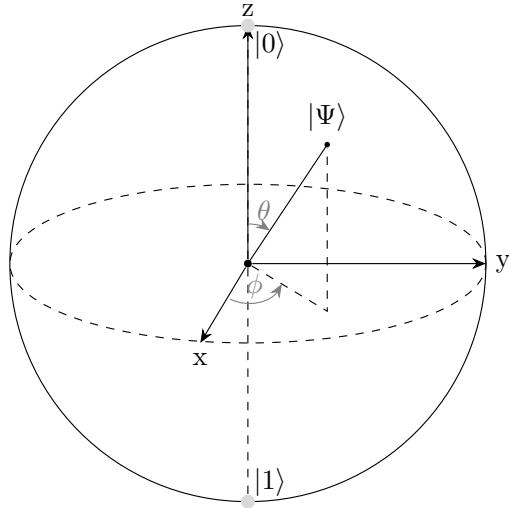


Figure 2.2: Bloch sphere representation of a single qubit quantum state.

is also popular. There will be four possible states, 2-bits of information could represent, only one at a time in classical mechanics. In contrast, in quantum mechanics, the state of superposition holds all possible states with some amplitudes at once. Thus an arbitrary two qubit state can be written in computational basis as

$$|\Psi\rangle = \alpha_{00}|00\rangle + \alpha_{01}|01\rangle + \alpha_{10}|10\rangle + \alpha_{11}|11\rangle, \quad (2.7)$$

where, α represents the amplitude, describing the probability of occurrence of the state attached to it as $|\alpha|^2$ upon measurement. Since the state is normalized, $\sum_{b \in \{0,1\}} |\alpha_b|^2 = 1$. One should notice that while there can be all four states in superposition, once the state gets measured, it collapses to give you one state, and this is probabilistic across many shots of measurement. This phenomenon is called state collapse.

2.2 Quantum properties

Until now, some new terms have been introduced already which a reader from classical computing background might find newer, yet sounding better. Quantum mechanics is interesting because of those fundamental properties, including superposition and entanglement, and they change how we perceive information processing.

2.2.1 Superposition

Quantum superposition is an intrinsic property of a qubit. While this may sound off, it is way easier to understand superposition with real life intuitive and analogous examples.

Let us consider that we have a two-sided coin with a head and a tail. We want to flip the coin and find whether it turns into a head or a tail. For convenience, we assume that turning into a head is a win and into a tail is a loss. While we flip the coin, and the coin is in air, it has equal probabilities of a win or a loss. We can only know if it is a win or a loss after seeing it stationary on the surface only. The state in which the coin existed when it was in the air is the state of superposition in a naive way. The basis states we talked about before are now the win or loss cases, and the amplitudes attached with the basis states give the probability of winning or losing.

For a single qubit, states $|+\rangle$ and $|-\rangle$ are the states of superposition and given in terms of computational basis states $|0\rangle$ and $|1\rangle$ as

$$\begin{aligned} |+\rangle &= \frac{|0\rangle + |1\rangle}{\sqrt{2}}, \\ |-\rangle &= \frac{|0\rangle - |1\rangle}{\sqrt{2}}. \end{aligned} \quad (2.8)$$

Measurement of these states in equation 2.8 give equal probabilities of landing on 0 or 1. Equation 2.7 represents the state in superposition between the basis states $|00\rangle$, $|01\rangle$, $|10\rangle$, and $|11\rangle$. A state represented by n qubits in superposition can explore the complex Hilbert space of dimension 2^n at the same time, and perform different quantum operations. Although we can process as many information at the same time, doing the measurement to read the results of computation becomes very tricky as it carries us to the classical picture of information, caused by state collapse into bits of 0 and 1.

2.2.2 Entanglement

Entanglement is a special case of superposition between two or more systems that describes how correlated the considered systems are. For two entangled systems, we cannot describe one system independently of the state of the other how far their separation is. This can be considered as a method of resource sharing between two systems in quantum information theory. Two arbitrary states, $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are said to be entangled if the entire state $|\Psi\rangle$ cannot be expressed as the product of individual states,

$$|\Psi\rangle \neq |\Psi_1\rangle \otimes |\Psi_2\rangle. \quad (2.9)$$

As long as a state $|\Psi\rangle$ can be represented as a product state, which is the product of two other states, and we can measure one participating state without affecting the state of the other. Let's see one of the Bell states, which is a maximally entangled state to illustrate the property of entanglement by some easier intuition. For two systems, A and B, each with a single qubit, the $|\Phi^+\rangle$ state in equation (2.10) is an entangled state.

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B). \quad (2.10)$$

For the $|\Phi^+\rangle$ state above, if we try to trace out the first system A, we won't know if it is $|0\rangle$ or $|1\rangle$ for A, thus we result in a maximally mixed state which we will discuss later in the text. Much simply, if we measure system A, and if the outcome of that measurement is 0(1), system B collapses to $|0\rangle(|1\rangle)$. This shows that the results of measurement on A have impact on the state B. This correlation is termed 'entanglement.'

However, if we have another system of two qubits $|\Phi'\rangle$ expressed as

$$|\Phi'\rangle = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |0\rangle_A \otimes |1\rangle_B), \quad (2.11)$$

we see that the system A is $|0\rangle$ in both terms, thus we can trace out this $|0\rangle_A$ state leaving the state $|+\rangle_B = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$ untouched. In simpler terms, while we measure A, if the outcome of measurement is 0, B is still in the $|+\rangle$ state. By 'tracing out,' we refer to removing the system A without touching the system B so that we can get a full picture of B.

In a two-qubit system, the Bell basis from equation 2.12 to 2.15 forms the orthonormal states, which are the maximally entangled states.

$$|\Phi^+\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \quad (2.12)$$

$$|\Phi^-\rangle = \frac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \quad (2.13)$$

$$|\Psi^+\rangle = \frac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \quad (2.14)$$

$$|\Psi^-\rangle = \frac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \quad (2.15)$$

Entanglement is the central concept in quantum information theory and the basis of quantum teleportation and quantum cryptography. The amount of entanglement contained in a system is measured as Von Neumann entropy for pure states.

For a state $|\Psi\rangle$ describing systems A and B, the reduced density matrix from the total state is given as

$$\rho_A = \text{Tr}_B(|\Psi\rangle\langle\Psi|). \quad (2.16)$$

The entanglement entropy or Von Neumann entropy, S is given as

$$S(\rho_A) = \text{Tr}(\rho_A \log_2 \rho_A). \quad (2.17)$$

Equivalently, $S(\rho_A) = -\sum_i \lambda_i \log_2 \lambda_i$ where $\{\lambda_i\}$ are the eigenvalues of ρ_A . For a two-qubit Bell state, the reduced density matrix of a subsystem A, $\rho_A = \frac{I}{2}$, which is a maximally mixed state and it gives $S = \log_2 2 = 1$. This shows that for a purely entangled system, the entanglement entropy of a subsystem is the maximum. It is necessary to understand that, for pure bipartite quantum states, the entanglement between subsystems is fully characterized by the von Neumann entropy of the reduced density matrix.

Highly entangled states are difficult to represent in classical computational methods like tensor networks, requiring larger dimension to describe the correlation. It is also difficult to simulate chemical systems where electrons are strongly correlated, which means the entanglement entropy between them is higher.

2.3 Quantum computation

2.3.1 Measurement

While classically measuring any physical quantity becomes very trivial most of the time, quantum mechanically, measurement is not that simple and requires a lot of care. Measurement in quantum computation is an operation that collapses the quantumness in a state and allows us to understand the resulting information in readable classical form.

Measurement operators act on the state space of the system being measured. If $\{M_m\}$ are the measurement operators and m represents the outcome of the corresponding measurement, for a state $|\Psi\rangle$, the probability of occurrence of the outcome m upon measurement is

$$p(m) = \langle \Psi | M_m^\dagger M_m | \Psi \rangle. \quad (2.18)$$

Since the state changes upon measurement, the resulting normalized state will be

$$|\Psi'\rangle = \frac{M_m |\Psi\rangle}{\langle \Psi | M_m^\dagger M_m | \Psi \rangle}. \quad (2.19)$$

A special class of measurement operators is the projective measurement. Projective measurement is described by an observable M on the state space of the system being measured. As we saw before, the probability of obtaining an outcome m upon measurement for a state $|\Psi\rangle$ is given by

$$p(m) = \langle \Psi | P_m | \Psi \rangle, \quad (2.20)$$

where P_m is obtained by spectral decomposition of the Hermitian operator M as $M = \sum_m m P_m$. If M_m in equation 2.18 are orthogonal projectors and $\sum_m M_m^\dagger M_m = I$, the general measurement expression reduces to projective measurements in equation 2.20.

These projective measurements can be used to obtain the average value of an observable, which is also called its expectation value. For an observable M , its expectation value is calculated as

$$\begin{aligned} E(M) &= \sum_m m p(m) \\ &= \sum_m m (\langle \Psi | P_m | \Psi \rangle) \\ &= \langle \Psi | \left(\sum_m m P_m \right) | \Psi \rangle = \langle \Psi | M | \Psi \rangle. \end{aligned} \quad (2.21)$$

If this observable is a Hamiltonian operator, the expectation value corresponding to this operator with respect to the state $|\Psi\rangle$ is the energy eigenvalue attached to that eigenstate.

Here, while we mention the term ‘observable,’ it means a Hermitian operator that represents some physical entity that can be measured. As already mentioned, an observable can be expressed as a spectral decomposition of the projector onto its eigenspace with some corresponding eigenvalue. While we want to measure any observable, we act the observable onto the system and measure the results on some basis.

2.3.2 Density operator

Density operator is a mathematical way of describing quantum system used mostly for those whose quantum state is not known completely. If i is the index that indexes the basis a quantum state $|\Psi\rangle$ can fall on with respective probabilities p_i , the density operator for this system is given as

$$\rho = \sum_i p_i |\Psi_i\rangle\langle\Psi_i|, \quad (2.22)$$

where $\{p_i, |\Psi_i\rangle\}$ is an ensemble of pure states.

For a closed system, if we apply a unitary operator U to this state $|\Psi_i\rangle$, the density operator ρ in equation 2.22 evolves as

$$\rho \rightarrow \sum_i p_i U |\Psi_i\rangle\langle\Psi_i| U^\dagger = U \rho U^\dagger. \quad (2.23)$$

A pure state we discussed here is a rank-1 projector. In simple terms, we can have complete knowledge of the system if it is a pure state. They can be expressed either by state-vector or as a density matrix. This state lies on the surface of the Bloch sphere.

In contrast, a mixed state is represented by a statistical mixture of pure states and can be described only by the density matrix. The mixed state lies inside the Bloch sphere and is a representation of reduced parts of the entangled states. They have non-zero entropy as opposed to the pure states which have zero entropy.

A maximally mixed state contains no information about the quantum system and the entropy is the maximum. While there are numerous ways we end up in a maximally mixed state, one was discussed in section 2.2. Tracing out one constituent system of a 2-qubit maximally entangled state leaves behind a single qubit maximally mixed state.

2.3.3 Quantum gates

Quantum gates are fundamental elements in circuit-based quantum computation for the implementation of unitary operators. Any arbitrary unitary operation can be implemented as a valid quantum gate.

(1) Single-qubit gates

The most fundamental gates in quantum computing are the Pauli- X , Y , Z , and I gates. All of them act on a single qubit state and change the state by unitary. In simple terms, acting these Pauli operators to a state $|\Psi\rangle$ is equivalent to rotating the state about X, Y or Z- axis on the Bloch sphere.

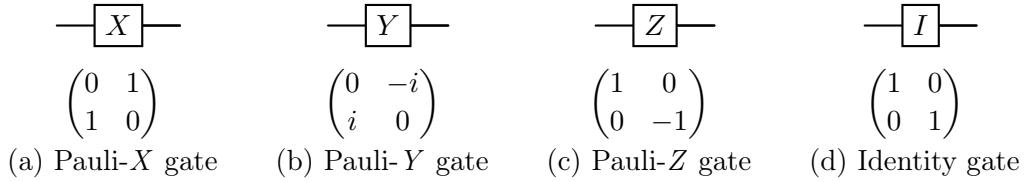


Figure 2.3: Single-qubit gates and their corresponding matrix representations.

Each of the Pauli gates rotate a state vector on the Bloch surface about the corresponding standard orthogonal axes. In addition to these, a very important gate exists which rotates the state vector about an axis halfway between the X and Z axes. This is Hadamard gate and it creates a superposition between two states in a qubit. The Hadamard gate is represented as shown in figure 2.4.

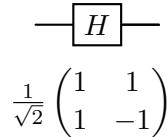


Figure 2.4: Hadamard gate and its matrix representation.

Thus the axis of rotation while using the Hadamard gate will be $\frac{X+Z}{\sqrt{2}}$. Hadamard gate maps $|0\rangle$ state into $\frac{|0\rangle+|1\rangle}{\sqrt{2}}$ and $|1\rangle$ state into $\frac{|0\rangle-|1\rangle}{\sqrt{2}}$, which are superposed states in the computational basis.

In general, for rotation by angle θ along the respective axis, the unitary rotation gates are expressed in general as

$$R_x(\theta) = e^{-i\theta X/2} = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) X = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -i \sin\left(\frac{\theta}{2}\right) \\ -i \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}, \quad (2.24)$$

$$R_y(\theta) = e^{-i\theta Y/2} = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) Y = \begin{pmatrix} \cos\left(\frac{\theta}{2}\right) & -\sin\left(\frac{\theta}{2}\right) \\ \sin\left(\frac{\theta}{2}\right) & \cos\left(\frac{\theta}{2}\right) \end{pmatrix}, \quad (2.25)$$

$$R_z(\theta) = e^{-i\theta Z/2} = \cos\left(\frac{\theta}{2}\right) I - i \sin\left(\frac{\theta}{2}\right) Z = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix}. \quad (2.26)$$

Pauli X , Y and Z gates are specific for an angle of $\theta = \pi$ radians. In addition to Pauli and Hadamard gates, single qubit gates like S and T gates are widely adapted for universal fault tolerant quantum computing. S and T gates are represented in matrices as

$$S = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} = \sqrt{Z}, \quad (2.27)$$

$$T = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{pmatrix} = \sqrt{S} = \sqrt{\sqrt{Z}}, \quad (2.28)$$

where,

$$Z = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\pi} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.29)$$

For $\theta = \pi$ in equation 2.26,

$$R_z(\pi) = \begin{pmatrix} e^{-i\frac{\pi}{2}} & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{pmatrix} = \begin{pmatrix} -i & 0 \\ 0 & i \end{pmatrix} = (-i) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = (-i)Z. \quad (2.30)$$

Since the global phases are unobservable, $R_z(\pi) \equiv Z$ up to a global phase. However, these are not identical matrices mathematically.

(2) Multi-qubit gates

Extending the case of single qubit gates, we have a widely adopted C-NOT gate which is a two-qubit gate. It is also called as Controlled-NOT gate with one qubit as control and the other as target. The gate flips the target qubit if and only if the control qubit is in state $|1\rangle$. The circuit representation of a CNOT gate is shown in figure 2.5. Equations in 2.31

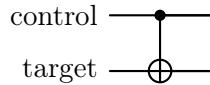


Figure 2.5: Controlled-NOT (CNOT) gate.

The quantum states on application of a CNOT gate translate as

$$\begin{aligned} |00\rangle &\rightarrow |00\rangle, \\ |01\rangle &\rightarrow |01\rangle, \\ |10\rangle &\rightarrow |11\rangle, \\ |11\rangle &\rightarrow |10\rangle. \end{aligned} \quad (2.31)$$

describe the action of a CNOT gate in a 2-qubit state. Based on these equations, we can

construct the matrix representation for the CNOT gate operation as

$$CNOT = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \quad (2.32)$$

We can understand this construction in many easy to understand ways. Since it is a 2-qubit gate, the space it extends is 2^n where n is 2. If we divide this 4×4 matrix into four compartments, the following resulting state is obtained due to NOT(Pauli X) in the target as

$$\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} \rightarrow \begin{pmatrix} I & 0 \\ 0 & X \end{pmatrix}. \quad (2.33)$$

Other than the Controlled - X gate, we have controlled- Z gate that introduces a phase flip to the target when the control qubit is in $|1\rangle$ state.

Besides, Toffoli-gate is also popular among multi-qubit gates which is a 3-qubit gate, with 2-qubits as control and one as target. It is represented as CCX, and applies NOT gate to the target qubit only if both the control qubits are in state $|1\rangle$. The circuit representation of a Toffoli gate is shown in figure 2.6.

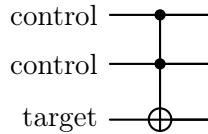


Figure 2.6: Toffoli gate.

2.4 A quick visit to molecular quantum mechanics

2.4.1 Wave function

A *wave function* in quantum mechanics is a complex valued function that contains all physically accessible information about the quantum system. It is represented as $|\Psi(x, t)\rangle$ which shows the dependence on x and t . The probability of finding a particle between x and $x + dx$ is given by the modulus square of the wavefunction, $|\Psi(x, t)|^2 dx$.

Since it is complex valued, it can be represented as $\Psi = \Re(\Psi) + i\Im(\Psi)$, where the imaginary part encodes the phase information of the wave function. A wave function is normalized as the total probability of finding a particle sums to 1, $\int |\Psi(x, t)|^2 dx = 1$. To find the wave function after time $t' = t + dt$, the wave function should be evolved according to the Schrödinger equation.

A wave function $\Psi(x, t)$ if separable can be written in terms of $\psi(x)\phi(t)$ where $\psi(x)$ depends only on x , and $\phi(t)$ depends only on t . However, not all functions are separable. In those cases, we can express them as the sum of separable functions.

2.4.2 Schrödinger equation

A generic Schrödinger equation in one dimension is given by

$$\frac{\partial \Psi}{\partial t} = \frac{i\hbar}{2m} \frac{\partial^2 \Psi}{\partial x^2} - \frac{iV\Psi}{\hbar}, \quad (2.34)$$

where V is the potential energy of the particle, which generally depends on x and t . Simplifying 2.34 lightly, we obtain

$$i\hbar \frac{\partial \Psi}{\partial t} = -\frac{\hbar^2}{2m} \cdot \frac{\partial^2 \Psi}{\partial x^2} + V\Psi. \quad (2.35)$$

We assume that V does not depend on time, thus we use $V(x)$ instead of $V(x, t)$. Solving equation 2.35,

$$\begin{aligned} i\hbar\psi \frac{d\phi}{dt} &= -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2}\phi + V\psi\phi, \\ \Rightarrow i\hbar \frac{1}{\phi} \frac{d\phi}{dt} &= -\frac{\hbar^2}{2m\psi} \frac{1}{dx^2} \frac{d^2\psi}{dx^2} + V. \end{aligned} \quad (2.36)$$

In the expression in equation 2.36, the expression on the left hand side has dependence on t (time) only, and on the right hand side has dependence on x (position) only. For any value of time and position, the LHS is equal to the RHS, which we take as a constant quantity. This quantity is the energy E contained in the system. Thus,

$$i\hbar \frac{1}{\phi} \frac{d\phi}{dt} = -\frac{\hbar^2}{2m\psi} \frac{1}{dx^2} \frac{d^2\psi}{dx^2} + V = E. \quad (2.37)$$

This gives rise to the time dependent and time independent versions of the Schrödinger equation.

(1) Time dependent Schrödinger equation

From equation 2.37, taking the time dependence part only,

$$i\hbar \frac{1}{\phi} \frac{d\phi}{dt} = E, \quad \Rightarrow \frac{d\phi}{dt} = -\frac{iE}{\hbar}\phi, \quad \Rightarrow \frac{d\phi}{\phi} = -\frac{iE}{\hbar}dt. \quad (2.38)$$

$$\begin{aligned} \text{Integrating both sides gives} \quad \int \frac{d\phi}{\phi} &= \int -\frac{iE}{\hbar}dt \\ \Rightarrow \ln |\phi| &= -\frac{iEt}{\hbar} + c. \end{aligned}$$

Exponentiating both sides, $e^{\ln |\phi|} = e^{-\frac{iEt}{\hbar} + c}$.

$$\text{Then, taking } e^c \text{ as constant, } \phi = e^{-\frac{iEt}{\hbar}}. \quad (2.39)$$

Equation 2.39 gives the time dependence of the separable solution of the Schrödinger equation and equation 2.38 shows the time dependent version of Schrödinger equation.

(2) Time independent Schrödinger equation

Similar to what we did for the time dependent case,

$$\begin{aligned} -\frac{\hbar^2}{2m} \cdot \frac{1}{\psi} \frac{d^2\psi}{dx^2} + V = E, \\ \Rightarrow -\frac{\hbar^2}{2m} \frac{d^2\psi}{dx^2} + V\psi = E\psi. \end{aligned} \quad (2.40)$$

Equation 2.40 is the time independent Schrödinger equation. If we define our Hamiltonian operator for a closed system as $\hat{H} = -\frac{\hbar^2}{2m} \frac{d^2}{dx^2} + V$, the equation 2.40 turns into an eigenvalue problem given in equation 2.41.

$$\hat{H}|\psi\rangle = E|\psi\rangle. \quad (2.41)$$

Here, $\psi(x)$ is not a single function, but can be a set of functions that describe a possible solution. Once we know $\psi(x)$, we can define $\Psi(x, t)$ as

$$\Psi(x, t) = \psi(x)e^{-\frac{iH}{\hbar}t}. \quad (2.42)$$

While $\psi(x)$ is a stationary state, we can use $\Psi(x, t)$ to form any possible wave function in the potential.

For the general solution to Schrödinger equation,

$$\Psi_n(x, t) = \psi_n(x)e^{-\frac{iE_n}{\hbar}t} = \sum_{n=1}^{\infty} c_n \psi_n(x)e^{-\frac{iE_n}{\hbar}t}, \quad (2.43)$$

where, $c_n = a_n + ib_n$ is the complex coefficient for a number of separable solutions. There might be a finite number of separable solutions in the general solution. We might start with a linear combination of two separable solutions, but as time evolves, it can go on changing. This wave function is oscillating as the time evolves. Throughout the range of this thesis, we consider the Hamiltonian to be time-independent Hamiltonian unless specified.

2.4.3 Spin operators

Spin is an intrinsic quantum mechanical property of a particle and does not have a clear classical analog. A moving charged particle creates a magnetic field according to classical electromagnetism. When the charged particles move around a circle, a magnetic field is produced as if a bar magnet is located at the center. In the case of an electron, which is a point particle, it cannot have anything else on its surface that could move around in an axis to generate magnetic fields. Schrödinger figured out the wobbling motion of electrons

called angular momentum, that kept the spin intact.

The famous Stern–Gerlach experiment opened a new paradigm on the study of spin particles [17]. A group of silver atoms with no resultant orbital angular momentum and no associated magnetic momentum pumped in the central pathway in the experiment got divided in half on either side of the large magnets used. Classically, they were expected to be randomly oriented at the end of the path, since an unpaired electron in a silver atom being in spherically symmetrical s-orbital, which did not happen. Magnetic moment in contrast to the expectation was present. This led us to describe that an electron has an intrinsic magnetic moment due to its intrinsic angular momentum, termed *spin*. It gave rise to two directions for explaining the electrons as spin-up and spin-down electrons with magnitude of $h/2\pi$. Pauli formulated a proper description of this quantum mechanical behavior using two dimensional matrices to represent the spin operator [18]. Trying to treat quantum mechanics and relativity together, Dirac figured out the effects of a spin with his popular Dirac equation [19].

Spin is described by a spin operator $\hat{\mathbf{S}} = (\hat{S}_x, \hat{S}_y, \hat{S}_z)$. The spin operator components obey the usual angular momentum commutation relations $[\hat{S}_j, \hat{S}_k] = i\epsilon_{jkl}\hat{S}_l$, where ϵ_{jkl} is the Levi-Civita symbol, $j, k, l = x, y, z$, and we work in units where $\hbar = 1$. The spin magnitude is described by the square of the spin operator $\hat{\mathbf{S}}^2$, which commutes with components of the spin operator $[\hat{\mathbf{S}}^2, \hat{S}_i] = 0$, for $i = x, y, z$. This leads to a common set of eigenvectors for $\hat{\mathbf{S}}^2$ and one of the spin components, usually taken to be \hat{S}_z . These eigenvectors $|s, m\rangle$ are labeled by the spin s and its projection along the z-direction m , and satisfy the following eigenvalue equations,

$$\hat{\mathbf{S}}^2|s, m\rangle = s(s+1)|s, m\rangle, \quad s = 0, \frac{1}{2}, 1, \frac{3}{2}, \dots \quad (2.44)$$

$$\hat{S}_z|s, m\rangle = m|s, m\rangle, \quad m = -s, -s+1, \dots, s. \quad (2.45)$$

The operator $\hat{\mathbf{S}}^2$ can be written in the following useful form,

$$\hat{\mathbf{S}}^2 = \frac{1}{2} \left(\hat{S}_+ \hat{S}_- + \hat{S}_- \hat{S}_+ \right) + \hat{S}_z^2, \quad (2.46)$$

where $\hat{S}_{\pm} = \hat{S}_x \pm i\hat{S}_y$ are the raising/lowering operators.

As an example, consider a single electron. Its spin component operators are given by the Pauli matrices $\hat{S}_i = \frac{1}{2}\hat{\sigma}_i$, for $i = x, y, z$. Since $s = \frac{1}{2}$ for an electron, there are two possible states,

$$|\alpha\rangle = \left| \frac{1}{2}, +\frac{1}{2} \right\rangle, \quad \text{and} \quad |\beta\rangle = \left| \frac{1}{2}, -\frac{1}{2} \right\rangle, \quad (2.47)$$

where we introduce α and β as a convenient notation for the spin-up and spin-down states of a single electron.

To make the concept of spin more clear, we describe the angular momentum using the cone model of spin as shown in figure 2.7. These figures depict the classical angular momentum we would have while a vector precesses around an axis (Z-axis here). We

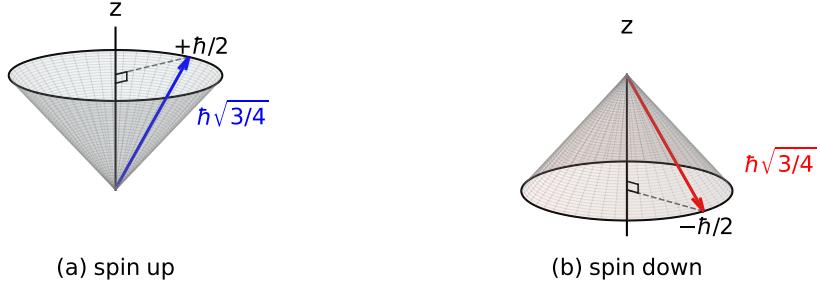


Figure 2.7: Cone diagrams to represent spin angular momentum. Blue and red arrows represent pictorial representation of a vector for spin angular momentum precessing around Z-axis. This is just a classical visualization for pedagogical purposes for understanding abstract quantum mechanical concepts of spin.

visualize this for an electron to make our study easier in a way that the projection of this precessing vector along the Z-axis is constant, $+\hbar/2$. This is the case of spin-up as shown by the figure on the left and mentioned in equation 2.47. As from above, since we have considered $\hbar = 1$, it gives us spin $+\frac{1}{2}$, and we call it α spin. Similarly the cone on the right gives us visualization of angular momentum with $-\frac{1}{2}$ projection on the Z-axis for β spin. If we consider a single electron and spinning up as in the figure on the left, $s = \frac{1}{2}$, $m = +\frac{\hbar}{2} \approx +\frac{1}{2}$. Length of spin will be $\hbar\sqrt{s(s+1)}$. Table 2.1 shows spin states and their corresponding quantum numbers and what they tell us about.

Let us consider two systems with spin operators $\hat{\mathbf{S}}_1$ and $\hat{\mathbf{S}}_2$. The total spin operator of the combined system is $\hat{\mathbf{S}} = \hat{\mathbf{S}}_1 + \hat{\mathbf{S}}_2$, with possible spin values S and the corresponding z-component M given by the addition theorem for angular momenta,

$$S = s_1 + s_2, s_1 + s_2 - 1, \dots, |s_1 - s_2|, \quad (2.48)$$

$$M = -S, -S + 1, \dots, S. \quad (2.49)$$

Here, s_1 and s_2 are the spin values of the individual subsystems, and we use capitalized S to emphasize the spin value for the combined system.

Those with $S = 1$ are called triplets and the one with $S = 0$ is called a singlet with no net spin. These are the four possible configurations two spins can be in.

Equation (2.46) can be generalized to the case of two subsystems,

$$\hat{\mathbf{S}}^2 = \hat{\mathbf{S}}_1^2 + \hat{\mathbf{S}}_2^2 + \hat{S}_{1+}\hat{S}_{2-} + \hat{S}_{1-}\hat{S}_{2+} + 2\hat{S}_{1z}\hat{S}_{2z}. \quad (2.50)$$

State	s_1	s_2	S	m_1	m_2	M	Description
$ \alpha\alpha\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	1	$+\frac{1}{2}$	$+\frac{1}{2}$	+1	Both spins up, length $\hbar\sqrt{2}$
$ \beta\beta\rangle$	$\frac{1}{2}$	$\frac{1}{2}$	1	$-\frac{1}{2}$	$-\frac{1}{2}$	-1	Both spins down, length $\hbar\sqrt{2}$
$ \alpha\beta\rangle$ (in phase)	$\frac{1}{2}$	$\frac{1}{2}$	1	$+\frac{1}{2}$	$-\frac{1}{2}$	0	Spins in phase, point same side in X-axis, rotates in XY-plane, length $\hbar\sqrt{2}$
$ \alpha\beta\rangle$ (out of phase)	$\frac{1}{2}$	$\frac{1}{2}$	0	$+\frac{1}{2}$	$-\frac{1}{2}$	0	Spins opposite in X and Y, total spin cancellation

Table 2.1: Spin states and their quantum numbers for two-spin systems.

The total N -electron spin eigenvectors are denoted by $|N, S, M\rangle$ and they satisfy similar eigenvalue equations to (2.44) and (2.45), provided the total spin value satisfies the constraint of (2.48).

2.4.4 Atomic and molecular orbital theory

An *orbital* is a mathematical function that describes the wavelike behavior or the probability distribution of an electron in an atomic or a molecular system. The magnitude square of this function gives the probability of finding the electron at a specific region around the nucleus. The concept of orbitals, and their shapes is derived from solving the electronic Schrödinger equation for hydrogen atom. While we don't have exact solution for atoms other than H atom, its solution is useful for describing the motion of individual electrons in other atoms. The solution yields a discrete set of energies that reflect the quantization of the motion of electrons in the Coulomb potential. The energy is given as

$$E_n = -\frac{m_e e^4}{32\pi^2 \epsilon_0^2 \hbar^2 n^2} = -\frac{13.6 eV}{n^2}, \quad ; n = 1, 2, 3, \dots \quad (2.51)$$

where m_e , e are the mass and charge of an electron and n is the principal quantum number that describes the energy of the state as well as the characteristic spatial extent of the probability distribution of the electron. For the hydrogen atom, the energy depends on n only, and the orbital is spherically symmetric.

As the number of electrons increases, we need more quantum numbers to describe the electronic structure in approximating many-body methods. The energies now depend on both n and the orbital angular momentum quantum number l . While n characterizes the radial extent and the overall energy scale of an electronic state as in H atom, l determines the shape of the orbital and the degree of penetration towards the nucleus. Orbitals with smaller l experience stronger effective nuclear attraction, leading to lower energies in comparison to the orbitals with higher l for the same n .

- For $l = 0$, s orbital which is spherically symmetric,
- For $l = 1$, p orbital with a single angular node (p_x, p_y, p_z),

- For $l = 2$, d orbital with two angular nodes ($d_{xy}, d_{yz}, d_{xz}, d_{x^2-y^2}, d_{z^2}$),
- For $l = 3$, f orbital with three angular nodes, seven total orientations.

Further, the label for orientation of the orbital is determined by the magnetic quantum number m_l , which takes integer values from $-l$ to $+l$. It specifies the projection of the orbital angular momentum along a chosen axis and is physically observable in the presence of magnetic field. Different m_l acquire different energies in the presence of external magnetic fields, leaving them degenerate in their absence. In the previous section, it is treated as m only in equations 2.44 and 2.45.

The orbitals we often use are the real-valued orbitals corresponding to the complex spherical harmonics. For $l = 1$, the complex spherical harmonics as indicated by m_l from Schrödinger equation are $Y_1^{-1}, Y_1^0, Y_1^{+1}$, where Y_1^0 is real and corresponds to p_z . The rest of Y_1^{-1} and Y_1^{+1} are complex and correspond to circular angular momentum around z . In chemistry, rather than using these complex terms, we use the real combinations as

$$p_x = \frac{1}{\sqrt{2}}(Y_1^1 + Y_1^{-1}), \quad p_y = \frac{1}{i\sqrt{2}}(Y_1^{+1} - Y_1^{-1}). \quad (2.52)$$

Another important thing to notice is that, while orbitals can be represented as different colors for different signs in them, these signs do not matter in general. It is because the probability density of finding a particle at location r is given as $\psi^*(r)\psi(r)$. These matter only when we go from one region of the orbital to the other. At nodes, the probability density of finding the particle is 0, with $\psi(r)$ being 0. Besides the angular nodes we talked about in connection with the value of l , there are other nodes too, which are the radial nodes, and these are formed at certain radius. To take an example, for a hydrogen atom, if we look into higher orbitals like 2s, 3s, 4s, it forms radial nodes moving from the center towards outside. For 2s orbital in a hydrogen atom, the probability of finding electrons goes to 0 at a certain radius, and becomes non-zero afterwards. When it becomes non-zero, it takes a different sign from the electron density in the center. The number of radial nodes is given by $n - l - 1$.

For convenience, it is assumed that electrons lie in orbitals in atoms, and filling the electrons in these orbitals follows Aufbau principle, which suggests to fill the lowest energy orbitals first. The filling up of orbitals follows the diagram shown in figure 2.8 and the energy in orbitals increases as we move bottom-left along those straight lines in the figure. The *electron configuration* lists electrons in each orbital. Each spatial orbital (s, p_x, p_y, p_z, \dots) can hold up to two electrons with spins α and β . For example, for a carbon atom with 6 electrons, the electron configuration is, $1s^2 2s^2 2p^2$.

This electron configuration doesn't tell about the spin (α, β) of the electrons in orbital. It is described by spin orbital configuration. For H atom, $1s$ in spin-orbital configuration becomes $1s\alpha$. For C atom whose electron configuration we had before, its spin orbital configuration follows Hund's rule, which suggests to put single electron in the orbitals with equal energy first, before pairing them up in each orbitals. Thus, the spin orbital configuration for C atom is, $1s\alpha 1s\beta 2s\alpha 2s\beta 2p_x\alpha 2p_y\alpha$ as shown in figure 2.9. These orbitals with equal energy, $2p_x, 2p_y, 2p_z$ in the figure are *degenerate orbitals*.

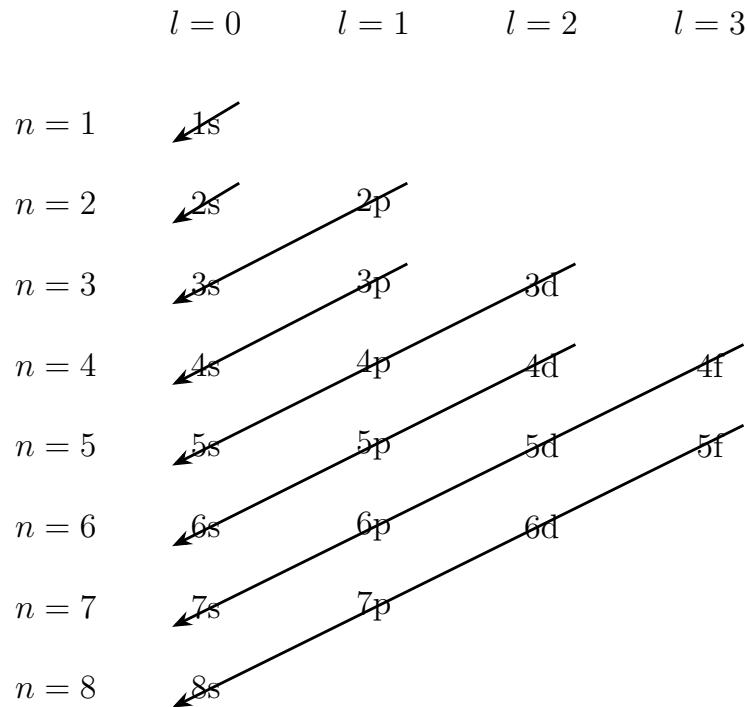


Figure 2.8: Orbital filling diagram according to Aufbau principle.

Now, when we go from atoms to molecules, these atomic orbitals combine to form molecular orbitals. Here, we talk about homo-nuclear, diatomic molecules like H_2 , N_2 only for simplicity.

Let us consider the molecular orbitals formed due to constructive and destructive interferences between the two s-orbitals, each from an atom as shown in figure 2.10. Here, Constructive interference merges the orbitals to one single orbital with the bond axis in the horizontal joining two nuclei. The bond axis is cylindrically symmetric, thus called σ orbitals. The subscript g in σ_g refers to gerade ('even' in German). It means that this orbital is even with respect to inversion symmetry. i.e., if we start from the middle of the orbital σ_g and go towards left/right or top/bottom, it is the same. In contrast to this, the destructive interference generates a node between the two nuclei. Destructively adding the two s-orbitals colored red in the figure can be thought of as adding two orbitals with opposite signs. So, the region they overlap turns into a node, which is a blank space as observed between two lobes in the anti-bonding orbital on the right. Looking at the bond axis now, it is still cylindrically symmetric, thus σ orbital. However, if we stay at the middle and go to the left(right), we see red(blue) phases, which shows anti-symmetry on inversion, thus ungerade orbital, represented with u in σ_u^* . '*' on this orbital representation denotes it is an anti-bonding orbital as there is a sign change while we go from one nucleus on the left

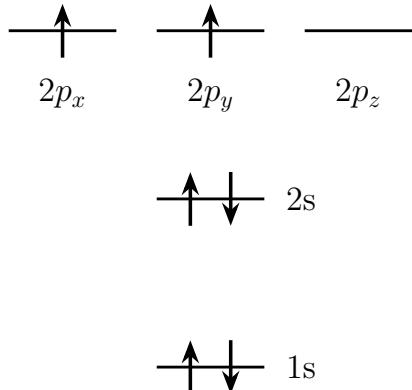


Figure 2.9: Carbon atom electron configuration in different orbitals following Hund's rule and Aufbau principle.

to the right or vice versa. On contrary, the upper figure has no sign change going from one nucleus to the other, thus it is a bonding orbital.

The bonding orbitals always have less energy and are more stable than the corresponding anti-bonding orbitals.

Similarly, when we look into two p_z orbitals, the formation of molecular orbitals takes place as shown in figure 2.11. In figure 2.11, on the top side, since we are adding the red lobe on the first orbital to the red lobe on the second orbital (due to the *minus* sign), constructive interference takes place and thus, a common red lobe is created at the center with the blue lobes surrounding it on two sides. Since the bond axis formed is symmetric cylindrically, it forms σ orbital. It is gerade as it is symmetric if we move up/down or left/right from the middle of the orbital (center of the red lobe). In addition, we confirm this is a bonding orbital as there is no sign change while we move from one nucleus to the other.

In contrast, destructive interference occurs between two p_z orbitals as in the lower configuration in figure 2.11, where the red lobe on the left orbital destructively interferes with the blue lobe on the right, causing a node at the center and reducing the size of either of these lobes. This creates an orbital which is still cylindrically symmetric along the bond axis, thus σ orbital. However, since sign change going towards the left/right from the center is not the same, it is not symmetric under inversion, thus ungerade. In addition, as the sign changes while moving from one nucleus to the other, it is an anti-bonding orbital.

Now we look into how the molecular orbitals are formed from p_x , p_x and p_y , p_y atomic orbitals in figure 2.12 and figure 2.13. When two p_x orbitals interfere constructively, the upper red and lower blue lobes form common red and blue lobes. In this case, the molecular orbital as seen in the upper right part of figure 2.12 is not symmetric cylindrically as there are red and blue lobes above and below the bond axis. Thus, since the lobes have sign

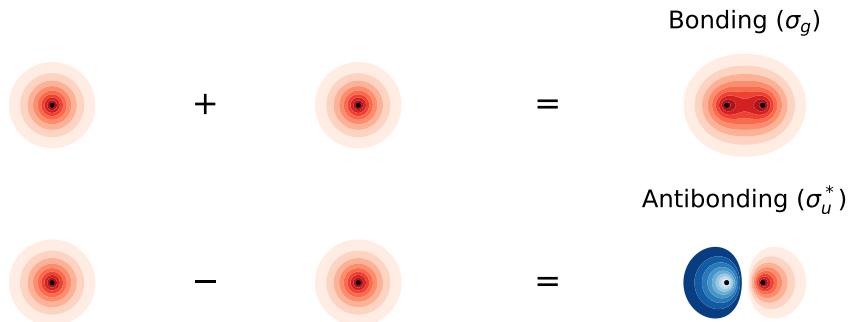


Figure 2.10: Molecular orbitals formed between s-s orbitals in two homo-nuclear atoms. Black dots at the center are nucleus in each atoms. The bonding orbitals are formed due to constructive interference between two s-orbitals, while the anti-bonding orbitals are the result of destructive interference.

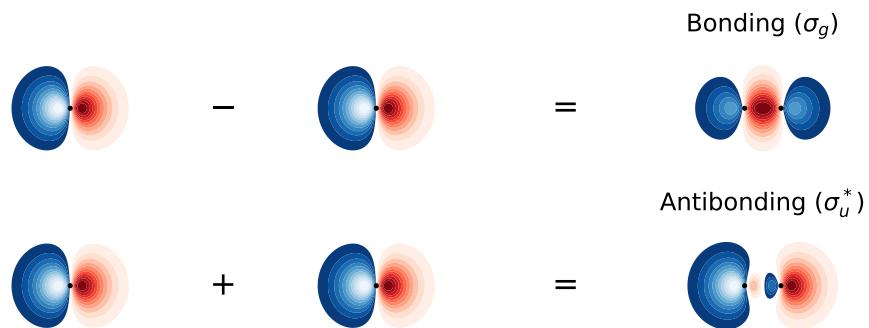


Figure 2.11: Molecular orbitals formed between p_z and p_z orbitals in two homo-nuclear atoms. Black dots at the center are nucleus in each atoms. The bonding orbitals are formed due to constructive interference between two p_z -orbitals, while the anti-bonding orbitals are the result of destructive interference.

change of π around the bond axis, this is a π orbital. Since there is a sign change if we go up/down the lobes from the center of the orbital, it is an ungerade. In addition to that, as we go from one nucleus and move towards the other by a slight shift from the bond axis, there is no sign change, thus forming a bonding orbital.

Similarly, as we subtract, destructive interference takes place. In this case as well, as we go across the bond axis there is a sign change by π , thus forming a π orbital. In addition, if we start from the center of the orbital and go towards the lobes in diagonal, it is the same in all lobes on diagonal, thus being gerade. Also, as we go from one nucleus and move towards the other, there is a sign change if we shift a little off the nuclear bond axis. Thus, this is an anti-bonding orbital.

The molecular orbital formation from p_y orbitals as shown in figure 2.13 follow the similar pattern as p_x orbitals.

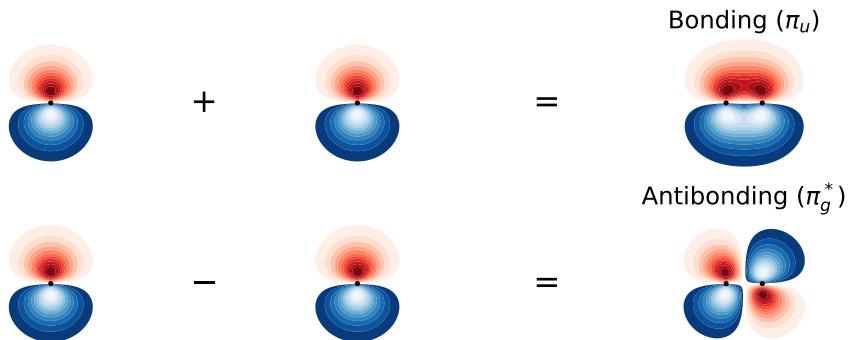


Figure 2.12: Molecular orbitals formed between p_x and p_x orbitals in two homo-nuclear atoms. Black dots at the center are nucleus in each atoms. The bonding orbitals are formed due to constructive interference between two p_x -orbitals, while the anti-bonding orbitals are the result of destructive interference.

Figure 2.14 shows how the molecular bonding and anti-bonding orbitals are formed from the atomic orbitals in two N atoms to form a N₂ molecule. The orbitals formation follows the pattern discussed in this section before. Following Aufbau principle, the electrons get filled up from the lowest energy orbitals first. The already filled up 1s orbitals are not shown in the picture as they have very less contribution. For N₂ molecule, filling up the orbitals in 2p gets completed within the low energy bonding orbitals, forming two π bonds and a single σ bond. The bond order of N₂ is 3 as there are six electrons in three bonding orbitals, and no electrons in anti-bonding orbitals since the bond order is given by the difference in the number of bonding and anti-bonding electrons divided by 2. This also describes the number of bonds between the atoms in a molecule. Setting up the levels of energy for molecular orbitals is not trivial. For molecules until O₂, $3\sigma_g$ orbital has higher energy than the $1\pi_u$ orbitals, which is not the case for O₂ and beyond.

While summarizing our visit to molecular orbitals theory, we have to still keep in mind some very basic foundations. Until now we have assumed that molecular orbitals are built from actual atomic orbitals, which is not entirely correct. The atomic orbitals we created were built from the knowledge of a hydrogen atom and are not exact for other atoms. This suggests that we can have more different atomic orbitals as atomic orbital basis functions and different molecular orbitals from them. The atomic orbitals we used were the minimal basis set used for picturing atoms.

2.4.5 Electronic structure problem

The electronic structure problem in quantum chemistry is the problem of solving the electronic wavefunctions with Born–Oppenheimer [20] approximation which approximates the behavior of the molecular Hamiltonian keeping the nuclei fixed. This is also called the eigenvalue problem as discussed in equation 2.41.

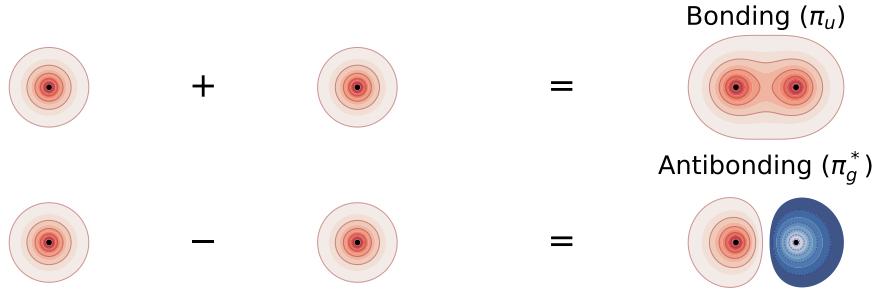


Figure 2.13: Molecular orbitals formed between p_y and p_y orbitals in two homo-nuclear atoms. Black dots at the center are nucleus in each atoms. The bonding orbitals are formed due to constructive interference between two p_y -orbitals, while the anti-bonding orbitals are the result of destructive interference. Here the visible red lobes have some blue pigment in them to describe that there are blue lobes on the other side of the plane. This can in overall be seen as bonding and anti-bonding orbitals formed from p_x orbitals and rotated by $\pi/2$.

(1) First quantization

In first quantization, where electrons are treated as individual particles with their own coordinates, rather than orbital occupation number, the electronic Hamiltonian \hat{H} is written as

$$\hat{H} = K_{nuc} + K_{elec} + V_{elec-nuc} + V_{elec-elec} + V_{nuc-nuc}, \quad (2.53)$$

where K_{nuc} and K_{elec} are the kinetic energies of nucleus and electron respectively, and $V_{elec-nuc}$, $V_{elec-elec}$ and $V_{nuc-nuc}$ are electron-nucleus attraction, electron-electron repulsion and nucleus-nucleus repulsion potential. This Hamiltonian is often known as the Coulomb Hamiltonian.

According to the Born-Oppenheimer approximation, we can treat all nuclei in a molecule as classical particles due to a very significant difference in their mass compared to the mass of electrons. Thus, the first quantized Hamiltonian in equation 2.53 for N electrons and M nuclei reduces to

$$\begin{aligned} \hat{H} &= K_{elec} + V_{elec-nuc} + V_{elec-elec} \\ &= -\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{i,k}^{N,M} \frac{Z_k}{|R_k - r_i|} + \sum_{i < j}^N \frac{1}{|r_i - r_j|} \end{aligned} \quad (2.54)$$

$$= \sum_i \hat{h}(i) + \sum_{i < j} v(i,j), \quad (2.55)$$

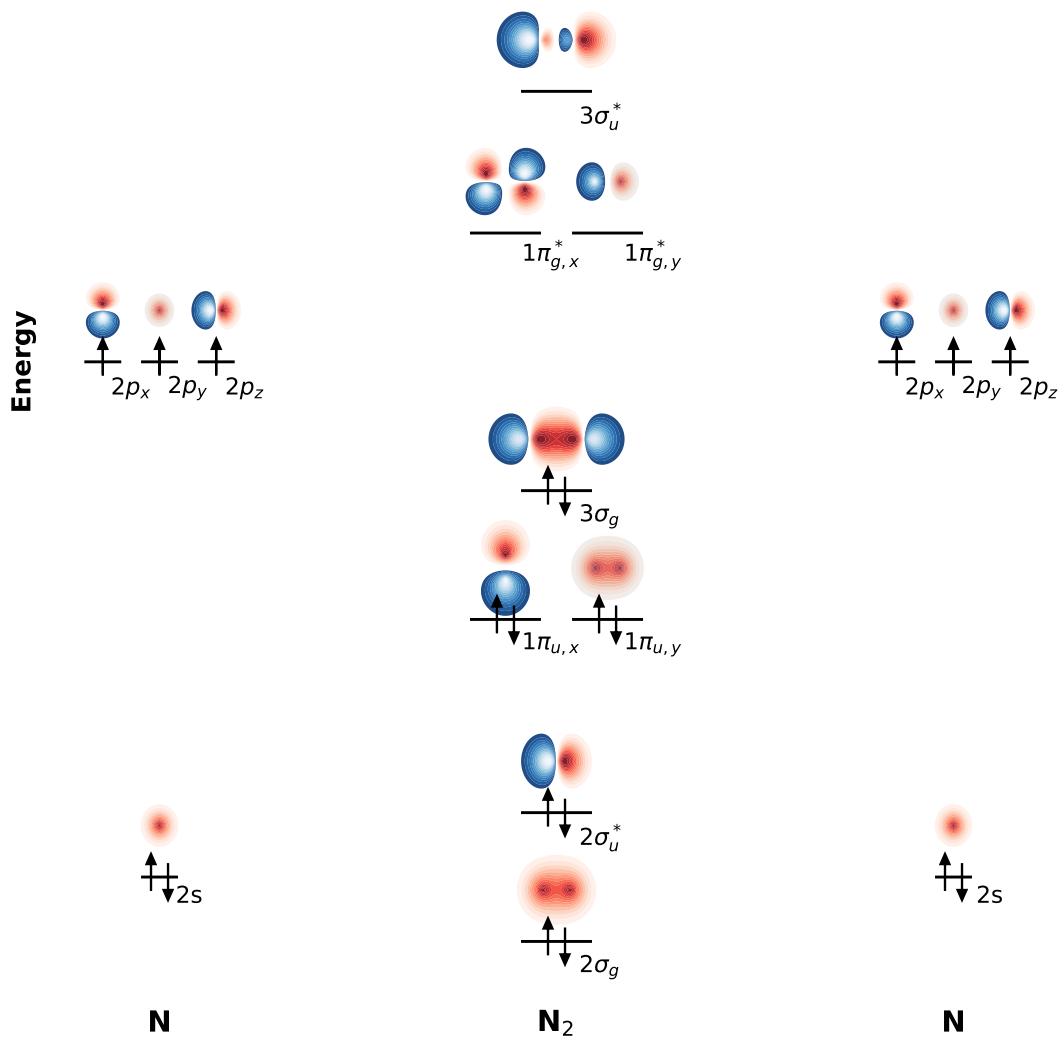


Figure 2.14: Formation of molecular orbitals in N_2 molecule using the atomic orbitals from two N atoms. The orbitals in the column in middle are molecular bonding and anti-bonding orbitals, with atomic orbitals on two sides.

where we have put 1 for $\frac{e^2}{4\pi\epsilon_0}$ in the second and third potential energy terms in equation 2.54 considering atomic units for the expression. ∇_i^2 is the Laplacian operator and is defined in Cartesian coordinates as $\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2}$, where (x_i, y_i, z_i) are the spatial coordinates of the i -th electron. Equation 2.55 is set for later use in the section where we discuss on Hartree–Fock method. Here, $\hat{h}(i)$ is one-electron operator for electron i and $v(i, j)$ is two electron operator for electrons i and j given as

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_i A} \quad \text{and} \quad v(i, j) = \frac{1}{r_{ij}}. \quad (2.56)$$

We represent the attraction from all nuclei here as \sum_A where A represents an atom.

(2) Second quantization

While the wave function in first quantization has to be manually antisymmetrized under electron exchange, another simpler and interesting form of quantization exists, called second quantization. Throughout the range of this thesis, the system we are considering is a fermionic system unless specified.

In second quantization, states with n particles representing spin orbitals can be either empty or filled up by a fermionic spinless particle [21]. The interactions of fermionic systems can be expressed using creation (\hat{a}_i^\dagger) and annihilation (\hat{a}_i) operators with $i \in \{0, \dots, n-1\}$ which enforce the antisymmetric property of fermions as

$$\{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}, \quad \{\hat{a}_i, \hat{a}_j\} = \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0. \quad (2.57)$$

This introduces a phase to the electronic basis state which has dependence on the occupancy of all orbitals with index less than i in the occupation number representation. This is a very important feature and will be useful when we derive qubit operators later in this section.

In the occupation number representation, the Fock space spanned by 2^n electronic basis states is $|f_{n-1} \dots f_0\rangle$ where $f_i \in \{0, 1\}$ is the occupation number of orbital i , and is 1 or 0 based on whether the orbital is occupied or not.

The creation and annihilation operators act on the Fock space basis vectors as

$$\hat{a}_i^\dagger |f_{n-1} \dots f_{i+1} 0 f_{i-1} \dots f_0\rangle = (-1)^{\sum_{p=0}^{i-1} f_p} |f_{n-1} \dots f_{i+1} 1 f_{i-1} \dots f_0\rangle, \quad (2.58)$$

$$\hat{a}_i^\dagger |f_{n-1} \dots f_{i+1} 1 f_{i-1} \dots f_0\rangle = 0, \quad (2.59)$$

$$\hat{a}_i |f_{n-1} \dots f_{i+1} 0 f_{i-1} \dots f_0\rangle = 0, \quad (2.60)$$

$$\hat{a}_i |f_{n-1} \dots f_{i+1} 1 f_{i-1} \dots f_0\rangle = (-1)^{\sum_{p=0}^{i-1} f_p} |f_{n-1} \dots f_{i+1} 0 f_{i-1} \dots f_0\rangle. \quad (2.61)$$

The molecular Hamiltonian being discussed can be expressed in the second quantized form as

$$\hat{H} = \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \frac{1}{2} \sum_{pqrs} h_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s, \quad (2.62)$$

where h_{pq} and h_{pqrs} are the one and two electron overlap integrals, respectively. One electron integral is obtained as

$$h_{pq} = -\frac{1}{2} \int \psi_p^*(r) \nabla^2 \psi_q(r) dr + \sum_k \int \psi_p^*(r) \frac{-Z_k}{|r - R_k|} \psi_q(r) dr, \quad (2.63)$$

where the first and second terms in equation 2.63 are kinetic energy and electron-nucleus attraction terms. Then, the two electron integral term is obtained as

$$h_{pqrs} = \int \int \psi_p^*(r_1) \psi_q^*(r_2) \frac{1}{|r_1 - r_2|} \psi_r(r_1) \psi_s(r_2) dr_1 dr_2, \quad (2.64)$$

where it describes the electron-electron repulsion between two electrons located at distance r_1 and r_2 .

(3) Fermion-to-qubit mapping

In order to simulate the system we are considering upon second quantization on a quantum computer, we have to translate it to the language the quantum computer understands. This Pauli-algebra model that a quantum computer understands is standard qubit model.

The Jordan–Wigner transformation (JWT) is the most widely used mapping method for fermionic system to map to qubits. Let us get a very clear understanding of how the Jordan–Wigner (JW) mapping works. We can map the electronic basis sets from the occupation number basis directly to the tensor product of qubit states in a quantum computer as

$$|f_{n-1} \dots f_0\rangle \rightarrow |q_{n-1}\rangle \otimes \dots \otimes |q_0\rangle, \quad (2.65)$$

where $f_i = q_i = \{0, 1\}$. Still, we have to map the creation and annihilation operators to some operators that can act on qubit with the similar action. Consider a single qubit creation operator \hat{C} and annihilation operator \hat{A} that can act on qubits as

$$\hat{C}|0\rangle = |1\rangle, \quad \hat{C}|1\rangle = 0, \quad \hat{A}|0\rangle = 0, \quad \hat{A}|1\rangle = |0\rangle.$$

These \hat{C} and \hat{A} operators can be expressed in matrices as,

$$\hat{C} = |1\rangle\langle 0| = \frac{1}{2}(X - iY), \quad \hat{A} = |0\rangle\langle 1| = \frac{1}{2}(X + iY).$$

Both \hat{C} and \hat{A} anticommute with Pauli-Z operator. So, to preserve the property of anticommutation relation of fermionic operators, we can apply Pauli-Z to the qubits with index below i and apply \hat{C} or \hat{A} based on the requirement to map \hat{a}_i^\dagger and \hat{a}_i respectively. This brings us to the final representation of fermionic operators in terms of qubit gates as

$$\hat{a}_i^\dagger \rightarrow Z^{\otimes i} \otimes \hat{C} \otimes 1^{\otimes(n-(i+1))}, \quad (2.66)$$

$$\hat{a}_i \rightarrow Z^{\otimes i} \otimes \hat{A} \otimes 1^{\otimes(n-(i+1))}. \quad (2.67)$$

Here, the operators \hat{a}_i^\dagger and \hat{a}_i are not unitary operators and we have to be conscious that we are not directly implementing the fermionic operators in quantum gates. Rather, after JWT, the fermionic Hamiltonian turns out to be a sum of Pauli strings, and the Paulis are unitary and Hermitian.

While this mapping seems very convincing, it needs $O(n)$ qubit operations for simulating n -qubit system. [22] proposed another mapping method called Bravyi–Kitaev transformation (BKT) with the requirement of $O(\log n)$ qubits for simulation. While *Bravyi et al.* have also proposed two methods (parity-basis and Bravyi–Kitaev basis), in the scope of this thesis I am going to talk only about the Bravyi–Kitaev basis only. Readers interested in the other method are recommended to visit [22], or [21] for easy to understand explanation. We have realized until now from the Jordan–Wigner transformation that two special information needed for fermionic operators are the occupation of the target orbital and the parity of the set of orbitals before the target orbital.

Bravyi–Kitaev transformation balances the locality of the occupation number as well as the parity information (which Jordan–Wigner lacked) to improve the efficiency of simulation. In this method, the qubit stores the parity of the set of occupation numbers corresponding to that set of orbitals. When i is even, the qubit on that index i stores the same orbital i . When i is odd, the qubit stores the set of adjacent orbitals with index less than i as partial sums $\sum_{p=k}^l f_p$ where k and l are determined by a different approach.

We map from the occupation number basis with f_i to the Bravyi–Kitaev basis b_j as

$$b_j = \sum_i [\beta_n]_{ji} f_i, \quad (2.68)$$

where β_n is given as,

$$\begin{aligned} \beta_{2^0} &= \beta_1^{-1} = [1], \\ \beta_{2^1} &= \beta_{2^0+1} = \left[\begin{array}{c|c} \beta_1 & 1 \\ \hline 0 & \beta_1 \end{array} \right] = \left[\begin{array}{c|c} 1 & 1 \\ \hline 0 & 1 \end{array} \right], \\ \beta_4 = \beta_{2^2} &= \beta_{2^1+1} = \left[\begin{array}{cc|cc} \beta_{2^1} & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \beta_{2^1} & \end{array} \right] = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{array} \right], \\ \beta_8 = \beta_{2^3} &= \beta_{2^2+1} = \left[\begin{array}{cccc|cccc} \beta_{2^2} & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right] = \left[\begin{array}{cc|cc|cc|cc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right]. \end{aligned} \quad (2.69)$$

Let us summarize the construction of $\beta_{2^{x+1}}$. We first divide the entire matrix into four quadrants; first, second, third and fourth. All entries in the third quadrant are 0s. In the first quadrant, all entries in the first row are 1s, and rest of the three rows are 0s. For the remaining two quadrants (second and fourth), we put β_{2^x} obtained from the previous step where we find $\beta_{2^x} = \beta_{2^{(x-1)+1}}$ using $\beta_{2^{x-1}}$. The base matrix is a 1×1 matrix as [1].

Then, we get the mapping from b_0 to b_7 as

$$\begin{bmatrix} b_7 \\ b_6 \\ b_5 \\ b_4 \\ b_3 \\ b_2 \\ b_1 \\ b_0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} f_7 \\ f_6 \\ f_5 \\ f_4 \\ f_3 \\ f_2 \\ f_1 \\ f_0 \end{bmatrix} = \begin{bmatrix} f_0 + f_1 + f_2 + f_3 + f_4 + f_5 + f_6 + f_7 \\ f_6 \\ f_4 + f_5 \\ f_4 \\ f_0 + f_1 + f_2 + f_3 \\ f_2 \\ f_0 + f_1 \\ f_0 \end{bmatrix}. \quad (2.70)$$

Thus, if we have a state $|11001100\rangle$ obtained as a result of occupation number mapping by Jordan–Wigner, the corresponding Bravyi–Kitaev mapping can be obtained as

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}. \quad (2.71)$$

Thus, the JW mapped state $|11001100\rangle$ corresponds to the state $|10001000\rangle$ with Bravyi–Kitaev mapping.

Before getting into a newer method of transformation, let's take a short look into the molecular Hamiltonian of H₂ [21], [23] with JWT in the STO-6G basis given as

$$\begin{aligned} \hat{H} = & c_0 1 + c_1 Z_0 + c_2 Z_1 + c_3 Z_2 + c_4 Z_0 Z_1 + c_4 Z_0 Z_2 + c_5 Z_1 Z_3 + c_6 X_0 Z_1 X_2 + c_6 Y_0 Z_1 Y_2 \\ & + c_7 Z_0 Z_1 Z_2 + c_4 Z_0 Z_2 Z_3 + c_3 Z_1 Z_2 Z_3 + c_6 X_0 Z_1 X_2 Z_3 + c_6 Y_0 Z_1 Y_2 Z_3 + c_7 Z_0 Z_1 Z_2 Z_3. \end{aligned} \quad (2.72)$$

While the Bravyi–Kitaev mapping also uses the same number of qubits for state representation as Jordan–Wigner, [24], [25] have proposed a method to reduce the qubit number based on symmetry. This is Symmetry-Conserving Bravyi–Kitaev Transformation (SCBKT) and it reduces the number of qubits for representing Hamiltonian by 2 qubits.

Let's see how SCBKT is obtained for a BK transformed Hamiltonian. In the Hamiltonian given in equation 2.72, the X and Y terms act only on qubits 0 (first) and 2 (third). Since the second and fourth qubits are always acted on by Z terms, we can remove these qubits from the Hamiltonian and replace these Paulis by their eigenvalues (± 1). (An worked out example of tapering off the qubits by this method is discussed in Appendix B of Bravyi *et al.* [25].)

Removing the qubits, we can rewrite the Hamiltonian \hat{H} as

$$\begin{aligned}\hat{H} = & c_0 1 + c_1 Z_0 + c_3 Z_2 + c_4 Z_0 Z_2 + c_6 X_0 X_2 + c_6 Y_0 Y_2 + c_7 Z_0 Z_2 \\ & + c_4 Z_0 Z_2 + c_6 X_0 X_2 + c_6 Y_0 Y_2 + c_7 Z_0 Z_2.\end{aligned}\quad (2.73)$$

Recalculating the coefficients and adding same terms, we get our final Hamiltonian as

$$\hat{H} = c'_0 1 + c'_1 Z_0 + c'_2 Z_1 + c'_3 Z_0 Z_1 + c'_4 X_0 X_1 + c'_5 Y_0 Y_1. \quad (2.74)$$

Here we saw that the qubits size needed to describe the H_2 molecule is 2 less than the one with full basis in STO-6G.

Similarly, we can obtain the qubit mapping for the SCBKT using JWT and BKT mapped qubit states. Lets look at the qubit state considered previously at equation 2.71,

$$\text{JWT- mapped: } |11001100\rangle \quad (2.75)$$

$$\text{BKT- mapped: } |10001000\rangle. \quad (2.76)$$

Considering $|f_0 f_1 f_2 f_3 f_4 f_5 f_6 f_7\rangle$, $|b_0 b_1 b_2 b_3 b_4 b_5 b_6 b_7\rangle$ for states $|1 1 0 0 1 1 0 0\rangle$ and $|1 0 0 0 1 0 0 0\rangle$ respectively, we see that the parity at corresponding indices of two states is even except for f_1/b_1 and f_4/b_4 . So, removing those bits with odd parity from the state, we obtain a state $|100100\rangle$, which is the final mapping for SCBKT [25].

2.4.6 Hartree–Fock theory

Hartree–Fock molecular orbital theory is used for the solution of the electronic Schrödinger equation. The electronic time-independent Schrödinger equation is given from 2.41 as

$$\hat{H}(r; R)\Psi(r; R) = E(R)\Psi(r; R). \quad (2.77)$$

This equation incorporating the Born–Oppenheimer approximation describes the wave function that depends on all the coordinates of the electrons r, and on all the coordinates of the nuclei R. It gives us electronic energy E which depends on where the nuclei are. This energy gives the potential energy surface for different coordinates of nuclei. We

can extract various properties from $\Psi(\mathbf{r}; \mathbf{R})$, $E(R)$ and their derivatives, like equilibrium geometry, vibrational frequencies, dipole moment, etc.

In Hartree–Fock method, the ground state wave function Ψ_{GS} can be approximated by Φ and expressed as

$$\Phi(x_1, x_2, \dots, x_N) = \hat{A}[\chi_1(x_1)\chi_2(x_2)\dots\chi_N(x_N)]. \quad (2.78)$$

Here, x_i in Φ are space-spin coordinates, which contain spatial orbital coordinate \mathbf{r} and spin ω as $x_i = (\mathbf{r}_i, \omega_i)$; $\mathbf{r}_i \in \mathbb{R}^3$ for (x -, y -, z -) and $\omega_i \in \{\uparrow, \downarrow\}$. The antisymmetrizer operator $\hat{A} = \frac{1}{\sqrt{N!}} \sum_{n=1}^N (-1)^{p_n} \hat{P}_n$, with sum over all $N!$ number of permutations P where p_n describing the number of pair inversions in \hat{P}_n . (By referring to antisymmetric property for fermions on exchange, we mean that the wave function changes its sign when two fermions swap their coordinates.)

The wave function Φ can be written as determinants, where the rows denote different electrons and the columns correspond to the spin orbital. This determinant structure is termed as Slater determinant and expressed in terms of spin orbitals $\chi_i(x_i)$ as

$$\Phi(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(x_1) & \chi_2(x_1) & \dots & \chi_N(x_1) \\ \chi_1(x_2) & \chi_2(x_2) & \dots & \chi_N(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \chi_1(x_N) & \chi_2(x_N) & \dots & \chi_N(x_N) \end{vmatrix}. \quad (2.79)$$

This is often written for convenience as $|\Phi\rangle = |\chi_1 \ \chi_2 \ \dots \ \chi_N\rangle$, where the spin orbitals χ are written as

$$\chi_i(x) = \psi_i(r)\alpha(\omega) \text{ or } \psi_i(r)\beta(\omega), \quad (2.80)$$

where $\alpha(\uparrow) = 1$, $\alpha(\downarrow) = 0$, $\beta(\uparrow) = 0$, $\beta(\downarrow) = 1$. An important property of these spin orbitals is orthonormality,

$$\langle \chi_i | \chi_j \rangle = \int \chi_i^*(x) \chi_j(x) dx = \delta_{ij}. \quad (2.81)$$

A single electron spin orbital χ describes an electron moving in the self-consistent mean-field of the nuclei and the average distribution of other electrons. Thus we say that Hartree–Fock wavefunction in equation 2.79 is a single Slater determinant constructed from one-electron mean-field spin orbitals. The Fock operator acts on one-electron orbital and gives back one-electron eigenvalue. One should be careful that though the Fock operator acts on a single electron orbital, that orbital already encodes the presence of other electrons through the mean field.

Lets recall the electronic Hamiltonian term derived previously in equation 2.55,

$$\hat{H} = \sum_i \hat{h}(i) + \sum_{i < j} v(i, j)$$

with one and two electron terms given as

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_A \frac{Z_A}{r_i A} \quad \text{and} \quad v(i, j) = \frac{1}{r_{ij}}.$$

Now, the electronic energy of the wave function is the average of the electronic Hamiltonian and given as

$$E(R) = \langle \Psi(\mathbf{r}; R) | \hat{H} | \Psi(\mathbf{r}; R) \rangle \quad (2.82)$$

$$= \int_{-\infty}^{\infty} \Psi^*(\mathbf{r}; R) \hat{H} \Psi(\mathbf{r}; R) d\mathbf{r}. \quad (2.83)$$

Then, using the Slater's rules as described in Chapter-2 of [26] and minimizing the expectation value of the exact Hamiltonian from equation 2.83, we get the relevant matrix elements and obtain Hartree–Fock energy as

$$E_{HF} = \sum_i \langle i | \hat{h} | i \rangle + \sum_{i < j} [ii|jj] - [ij|ji], \quad (2.84)$$

where $\langle i | \hat{h} | i \rangle$ is a one electron integral,

$$\langle i | \hat{h} | j \rangle = \int \chi_i^*(\mathbf{x}_1) \hat{h}(\mathbf{r}_1) \chi_j(\mathbf{x}_1) d\mathbf{x}_1. \quad (2.85)$$

Here \mathbf{x} contains (x,y,z,ω) components, while \mathbf{r} contains (x,y,z) only.

This Hartree–Fock energy obtained in equation 2.84 is not equivalent to summing up the eigenvalues of the fock operator on one-electron orbitals, since the earlier one is obtained by variational methods by minimizing the expectation value of the exact many-electron Hamiltonian evaluated on the resulting determinant formed from single-electron orbitals.

Similarly, $[ii|jj]$ and $[ij|ji]$ are two-electron integral terms. Their general form is given as

$$[ij|kl] = \int \chi_i^*(\mathbf{x}_1) \chi_j(\mathbf{x}_1) \frac{1}{r_{12}} \chi_k^*(\mathbf{x}_2) \chi_l(\mathbf{x}_2) d\mathbf{x}_1 d\mathbf{x}_2. \quad (2.86)$$

While the one-electron integral is four-dimensional, the two-electron integral term is eight-dimensional due to the presence of two \mathbf{x} terms in the integral. In equation 2.86, r_{12} is the distance between electron-1 and electron-2. We see that one electron at \mathbf{x}_1 is in two different orbitals, χ_i and χ_j , which is not intuitive physically. This is just a mathematical idea and does not actually make sense in the real world.

Now, let's look at the one-electron term in equation 2.84,

$$\langle i | \hat{h} | i \rangle = \int \chi_i^*(\mathbf{x}_1) \hat{h}(\mathbf{r}_1) \chi_i(\mathbf{x}_1) d\mathbf{x}_1. \quad (2.87)$$

This term looks like the term we would see if we wanted to calculate the expectation value of operator \hat{h} for one occupied orbital i . Thus, it describes the average kinetic energy of

electron 1 plus the potential of attraction to all other nuclei, since \hat{h} consists of kinetic energy and potential energy terms. Getting the sum over all occupied orbitals gives us our one-electron contribution term for the energy.

Similarly, lets look at the two-electron terms in equation 2.84 separately as

$$[ii|jj] = \int d\mathbf{x}_1 \int \chi_i^*(\mathbf{x}_1) \chi_i(\mathbf{x}_1) \frac{1}{r_{12}} \chi_j^*(\mathbf{x}_2) \chi_j(\mathbf{x}_2) d\mathbf{x}_2. \quad (2.88)$$

The two-electron term in equation 2.88 arise due to electron-electron repulsion and is called Coulomb integral. The terms $\chi_i^*(\mathbf{x}_1) \chi_i(\mathbf{x}_1)$ ($\chi_j^*(\mathbf{x}_2) \chi_j(\mathbf{x}_2)$) give the probability of finding an electron 1(2) at location $\mathbf{x}_1(\mathbf{x}_2)$ if the electron was in orbital $\chi_i(\chi_j)$. Since we do not know where the electrons are, we have to integrate over all possible places they might be in. The equation 2.88 is thus considered as a generalized form of the Coulomb law as we know exactly where the electrons lie in the case of normal Coulomb law with which we are familiar. If the positions of electrons were known , the probability terms in the left and right of $\frac{1}{r_{12}}$ would equal to 1, and thus the two-integral term would be Coulomb repulsion between electron 1 in orbital i and electron 2 in orbital j .

Now looking at the second term in two-electron term in equation 2.88

$$[ij|ji] = \int d\mathbf{x}_1 \int \chi_i^*(\mathbf{x}_1) \chi_j(\mathbf{x}_1) \frac{1}{r_{12}} \chi_j^*(\mathbf{x}_2) \chi_i(\mathbf{x}_2) d\mathbf{x}_2. \quad (2.89)$$

This term is called the exchange integral and looks almost the same as the Coulomb integral except that the indices i and j in $\chi_i(\mathbf{x}_1)$ and $\chi_j(\mathbf{x}_2)$ are exchanged from the Coulomb integral in equation 2.88.

The integrals in equation 2.87 can be separated into integral over spin coordinate (ω) and spatial coordinates \mathbf{r} which contains x, y, z as

$$\begin{aligned} \langle i|\hat{h}|j\rangle &= \int \phi_i^*(\mathbf{r}) \sigma_i^*(\omega) \hat{h}(\mathbf{r}) \phi_j(\mathbf{r}) \sigma_j(\omega) d\mathbf{r} d\omega \\ &= \int \phi_i^*(\mathbf{r}) \hat{h}(\mathbf{r}) \phi_j(\mathbf{r}) d\mathbf{r} \int \sigma_i^*(\omega) \sigma_j(\omega) d\omega \\ &= \left[\int \sigma_i^*(\omega) \sigma_j(\omega) d\omega \right] \cdot (i|\hat{h}|j), \end{aligned} \quad (2.90)$$

where $(i|\hat{h}|j)$ is the same as $\langle i|\hat{h}|j\rangle$ except that now the integral is over spatial coordinates only and not the spin part and $\phi(\mathbf{r})$ and $\sigma(\omega)$ are spatial and spin parts respectively corresponding to the spin orbital $\chi(\mathbf{x})$.

Similarly, if we separate the spatial and spin parts in two-electron integral part,

$$\begin{aligned} [ij|kl] &= \int d\mathbf{r}_1 \int d\omega_1 \int d\mathbf{r}_2 \int d\omega_2 \phi_i^*(\mathbf{r}_1) \sigma_i^*(\omega_1) \phi_j(\mathbf{r}_1) \sigma_j(\omega_1) \frac{1}{r_{12}} \phi_k^*(\mathbf{r}_2) \sigma_k^*(\omega_2) \phi_l(\mathbf{r}_2) \sigma_l(\omega_2) \\ &= \int \sigma_i^*(\omega_1) \sigma_j(\omega_1) d\omega_1 \int \sigma_k^*(\omega_2) \sigma_l(\omega_2) d\omega_2 \cdot \int d\mathbf{r}_1 \int \phi_i^*(\mathbf{r}_1) \phi_j(\mathbf{r}_1) \frac{1}{r_{12}} \phi_k^*(\mathbf{r}_2) \phi_l(\mathbf{r}_2) d\mathbf{r}_2 \\ &= \left[\int \sigma_i^*(\omega_1) \sigma_j(\omega_1) d\omega_1 \int \sigma_k^*(\omega_2) \sigma_l(\omega_2) d\omega_2 \right] \cdot (ij|kl), \end{aligned} \quad (2.91)$$

where $\langle ij|kl \rangle$ is the same as $\langle ij|kl \rangle$ except that now the integral is over spatial coordinates only and not the spin part and $\phi(\mathbf{r})$ and $\sigma(\omega)$.

Separating the spatial and spin part helps as the integrals over spin coordinates are either 0 or 1. While σ is either α or β ,

$$\int \alpha^*(\omega)\alpha(\omega)d\omega = 1, \quad \int \beta^*(\omega)\beta(\omega)d\omega = 1, \quad \int \alpha^*(\omega)\beta(\omega)d\omega = 0, \quad \int \beta^*(\omega)\alpha(\omega)d\omega = 0.$$

Thus,

$$\begin{aligned} \langle i|\hat{h}|j \rangle &= \left[\int \sigma_i^*(\omega)\sigma_j(\omega)d\omega \right] \cdot (i|\hat{h}|j) \\ &= \begin{cases} (i|\hat{h}|j) & \text{when } \sigma_i = \sigma_j \\ 0 & \text{when } \sigma_i \neq \sigma_j \end{cases}. \end{aligned} \quad (2.92)$$

We understand that the one-electron integrals remain only when both spin orbitals have the same spin.

Similarly,

$$[ij|kl] = \begin{cases} (ij|kl) & \text{when } \sigma_i = \sigma_j \text{ and } \sigma_k = \sigma_l \\ 0 & \text{when } \sigma_i \neq \sigma_j \text{ or } \sigma_k \neq \sigma_l. \end{cases} \quad (2.93)$$

From here, we observe that if the first two and last two indices of the spin orbital are the same spin, the spatial terms remain. In any of the cases when the indices differ, the two-electron integrals do not live.

From the general solution of the composing integrals, we obtain each of the integrals in Hartree–Fock as

$$\begin{aligned} \langle i|\hat{h}|i \rangle &= (i|\hat{h}|i) && \text{spatial integral term always survives} \\ [ii|jj] &= (ii|jj) && \text{spatial integral term always survives} \\ [ij|ji] &= (ij|ji) && \text{spatial integral term survives only if } \sigma_i = \sigma_j. \end{aligned} \quad (2.94)$$

To strengthen our understanding, lets take an example where two electrons with opposite spins (α) and (β) sit on the same spatial orbital $1s$. Following the notation used in [26], we denote electrons with α and β spins as 1 and $\bar{1}$ respectively. Also, while getting the integrals, we have already proved that after spin integration to 0 or 1, only the spatial terms survive, and as the spatial orbital is the same for this example we are taking, the spatial integrals have to be the same too. Hence, bars ($\bar{}$) above the electrons are not required. Thus, the Hartree–Fock energy will be

$$\begin{aligned} E_{HF} &= \sum_i \langle i|\hat{h}|i \rangle + \sum_{i < j} [ii|jj] - [ij|ji] \\ &= \langle 1|\hat{h}|1 \rangle + \langle \bar{1}|\hat{h}|\bar{1} \rangle + [11|\bar{1}\bar{1}] - [1\bar{1}|\bar{1}1] \\ &= (1|\hat{h}|1) + (\bar{1}|\hat{h}|\bar{1}) + (11|11) \\ &= 2(1|\hat{h}|1) + (11|11). \end{aligned} \quad (2.95)$$

Here, $[1\bar{1}|\bar{1}1] = 0$ and $[11|\bar{1}\bar{1}]$ will be equal to $(11|11)$ since α and β pair with the same spatial part.

This computational method in which the spin orbitals always come in the pairs of (α, β) and stay in the same spatial orbital is called Restricted Hartree–Fock (RHF). The spin orbitals in the RHF method are given as

$$\begin{aligned}\chi_n(\mathbf{x}) &= \phi_n(\mathbf{r})\alpha(\omega) \\ \chi_{\bar{n}}(\mathbf{x}) &= \phi_{\bar{n}}(\mathbf{r})\beta(\omega).\end{aligned}\quad (2.96)$$

RHF is most commonly used for closed shell molecules when all electrons are paired. For the case of open-shell molecules where not all the electrons are paired, we can get a lower energy solution when we do not restrict the pair of α and β to occupy the same spatial orbitals, i.e., allow the spatial part for α and β spin to be different mathematically. This method of computation is called Unrestricted Hartree–Fock (UHF) and the spin orbitals in UHF are given as

$$\begin{aligned}\chi_n(\mathbf{x}) &= \phi_n(\mathbf{r})\alpha(\omega) \\ \chi_{\bar{n}}(\mathbf{x}) &= \phi_{\bar{n}}(\mathbf{r})\beta(\omega).\end{aligned}\quad (2.97)$$

This introduces spin contamination where the spin of different multiplicities mix.

This leads to an easier description of Hartree–Fock energy from a bigger picture as,

$$E_{HF} = \sum_i h_{ii} + \sum_{i < j} J_{ij} + (-K_{ij}), \quad (2.98)$$

where h_{ii} is one-electron term, J_{ij} is Coulomb repulsion from each unique pair of electrons and $-K_{ij}$ is the exchange term from each unique pair of same spin.

UHF is a case of symmetry breaking. RHF is based on the foundational condition that two opposite spin electron pairs live in the same spatial orbital and the spatial orbital is always localized over both atoms. However, in UHF, the electron pairs live in different spatial orbitals with same spin. Having same spin breaks the spin symmetry as two spins in two electrons are expected to have opposite spin from one another. Normally only spin symmetry is broken in the case of UHF. However, sometimes both spin and spatial symmetry could be broken too. Spin symmetry breaking is easy to explain as it is having two spin orbitals in different spatial orbitals. However, spatial symmetry breaking is guaranteed when the electrons are localized towards the left and right atoms (during bond dissociation, in strongly correlated systems). By symmetry breaking, we are referring to both spin and spatial symmetry breaking in the range of this thesis.

Connection to Hund's Rules

Hartree–Fock theory can also be used to understand Hund's rules. While this concept is not necessary for this thesis, I learned this during my research and was very excited to know about it. I had studied Hund's rule during my high-school and had no idea why it holds true.

Consider two different configuration of two electrons in p-orbital as shown in figure 2.15 where (a) has unpaired electrons in different spatial orbitals, each with different spin term and (b) has a pair of opposite spin electrons in the same spatial orbital. These form an example of UHF and RHF configuration. According to Hund's rules, for a given electron configuration, state with higher spin is more stable than the state with lower spin. So, following his rules, the configuration in (a) is more stable as it has higher net spin than configuration in (b). Let's discuss why it is true.



Figure 2.15: Comparison of two orbital occupation configurations across the three p -orbitals.

For configuration (a), $E_{HF_a} = h_{xx} + h_{yy} + J_{xy} - K_{xy}$. Similarly, for configuration (b), $E_{HF_b} = 2h_{xx} + J_{xx}$. (b) has no K term as the exchange term is 0 when both electrons spin in different directions. So, if we compare the energies E_{HF_a} and E_{HF_b} for the configuration, R_{HF_a} is lower than E_{HF_b} since the exchange term is subtracted from almost similar energy in the case of E_{HF_a} and $K \geq 0$. Thus, the energy of the configuration on the left is lower than the energy of the configuration on the right, thus (a) is more stable than (b).

When we have discussed describing the wave function with Slater determinants, a single Slater determinant is enough to describe that wave function which has a single dominant electronic configuration. Lets look at a singlet ground state for H_2 molecule, which is an electronic configuration with zero spin multiplicity as in the figure, 2.16. In the singlet

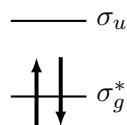


Figure 2.16: H_2 molecule singlet ground state configuration showing antiparallel spins in the bonding σ_u orbital.

configuration of figure 2.16, the electrons in the ground state are stable in the bonding orbital σ_g , thus this state has a single dominant configuration, and can be described by

a single Slater determinant. Let's see at the open-shell singlet state as shown in the figure 2.17. Here, we see two different configuration of electrons in the orbitals. Both on

$$|\Psi_{S=0,M=0}\rangle = \frac{1}{\sqrt{2}} \left(\begin{array}{c} \uparrow \\ \downarrow \end{array} \right) - \frac{1}{\sqrt{2}} \left(\begin{array}{c} \downarrow \\ \uparrow \end{array} \right)$$

Figure 2.17: Open-shell singlet state as a superposition of spin configurations.

the left and the right have the same energy, and we describe the open-shell singlet state as a combination. Since this is a singlet, it has $S = 0$ and $M = 0$. If we had a plus, ‘+’

$$|\Psi_{S=1,M=0}\rangle = \frac{1}{\sqrt{2}} \left(\begin{array}{c} \uparrow \\ \downarrow \end{array} \right) + \frac{1}{\sqrt{2}} \left(\begin{array}{c} \downarrow \\ \uparrow \end{array} \right)$$

Figure 2.18: Low-spin triplet state as a superposition of spin configurations.

in between the two configurations as in figure 2.18, it would be a low-spin triplet state with $S = 1$ and $M = 0$. Just to let the reader know, H_2 molecule has two other triplet states with both spins up ($S = 1, M = 1$) and both spins down ($S = 1, M = -1$). But these later two triplets can be described by a single Slater determinant as they have only one dominant configuration each. However for the triplet with $S = 1$ and $M = 0$ and the open shell singlet in figure 2.17, we need two Slater determinants to describe either of them. The configuration on the left in figure 2.17 and the configuration on the right need separate Slater determinants to describe them. Thus, we need two in this case.

On the other hand, if we consider only one configuration (let the first configuration) of both singlet and triplet states in figure 2.17 and 2.18, the resulting state looks like as shown in figure 2.19. This state has no spin symmetry and is not the eigen function of the

$$\begin{array}{c} \uparrow \\ \downarrow \end{array} = \frac{1}{\sqrt{2}} (|\Psi_{S=0,M=1}\rangle + |\Psi_{S=1,M=0}\rangle)$$

Figure 2.19: Superposition of singlet and triplet states.

$\hat{\mathbf{S}}^2$ operator. We talk about this kind of state later in the course of this thesis.

The requirement of multiple Slater determinants to describe even the simplest two-electron triplet state highlights the intrinsic multi-determinantal nature of certain electronic states. Multi-configuration self-consistent field methods provide a systematic framework to address this issue by explicitly constructing and optimizing a linear combination of

determinants. However, such approaches rapidly become complex as the number of relevant configurations increases. Moreover, once the electronic structure becomes strongly correlated, the notion of canonical Hartree–Fock orbitals and their associated orbital energies loses physical significance. These considerations motivate a complementary perspective that does not rely on the explicit enumeration of Slater determinants, but instead focuses on quantities that remain well defined for arbitrary many-electron wavefunctions. In this context, the one-body reduced density matrix provides a natural starting point.

(1) One-body reduced density matrix

The one-body reduced density matrix (1-RDM) provides a compact and basis-independent description of all single-particle properties of a many-electron wavefunction. It is defined by tracing out $N - 1$ electrons from the full many-body density operator. Unlike the Fock operator, which is meaningful only within the Hartree–Fock approximation, the 1-RDM is well defined for arbitrary correlated wavefunctions, including multi-determinantal and symmetry-broken states. As such, it serves as a natural object for analyzing electronic structure beyond the mean-field level, encoding information about electron populations, coherence between orbitals, and the degree of electronic correlation present in the system.

Natural orbitals are defined as the eigenfunctions of the one-body reduced density matrix, with the corresponding eigenvalues referred to as occupation numbers. In weakly correlated systems, natural orbitals closely resemble Hartree–Fock canonical orbitals and exhibit occupation numbers close to 0 or 2. In contrast, in strongly correlated regimes such as bond dissociation, several natural orbitals acquire fractional occupations, directly reflecting the breakdown of a single-determinant description. Since natural orbitals do not diagonalize the Fock operator, they are not associated with orbital energies; instead, their occupation numbers provide a physically meaningful characterization of correlation effects.

While we use natural orbitals as the basis functions for computing molecular orbitals in this work, figures including figure 3.2 still use the representations for molecular orbitals created with atomic orbitals.

2.4.7 Density functional theory

While the wavefunction based methods rely on many electron wavefunction as a linear combination of Slater determinants, there are other approaches that reformulate the electronic structure problem in terms of the electron density, giving rise to density functional theory (DFT). DFT avoids the explicit many-body wavefunctions whose complexity scale exponentially with system size. In general, DFT is implemented through the Kohn-Sham (KS) scheme that maps the interacting many-electron problem to an auxiliary non-interacting system which reproduces the same electron density. This approach enables efficient simulations of large systems and incorporates electron correlation at relatively low computational cost. Since the exact exchange-correlation functional in DFT is unknown, it uses approximate density functionals, and struggles in capturing strong

correlation properly. DFT however can be very helpful complementary subroutines for quantum computational chemistry tasks. DFT methods can also be taken as substitutes of wavefunction based approaches for computational efficiency.

2.4.8 Strongly correlated systems

(1) Open-shell systems

Open shell systems are those molecular systems carrying unpaired electrons. When the molecule has more than one unpaired electron, electronic states of different spin multiplicities appear in the low energy region, and they are representative of strongly correlated systems. Open shell systems play an important role in chemistry and materials science as potential candidates for magnetic materials [27] and nonlinear optical materials [28]. In addition, multinuclear transition metal complexes containing unpaired d-electrons are involved in a number of catalytic processes and enzymes, such as Mn₄CaO₅ cluster in the photosystem II [29] and FeMoco cluster in nitrogenase [30]. It is important to note that open shell electronic structure appears even in small, simple molecules such as H₂ when covalent bond dissociation takes place.

Let us consider an example of H₂ molecule to discuss about open-shell system, bond-dissociation and Hartree–Fock methods for the corresponding state of molecule.

For the H₂ molecule, the lowest energy molecular orbital is the gerade combination. A wave function obeys gerade symmetry if it remains unchanged on inversion through the center of symmetry, i.e., at the midpoint of the bonding axis. The orbital is given as

$$\sigma_g(\mathbf{r}) = \frac{1}{\sqrt{2(1+S)}}(\phi_a(\mathbf{r}) + \phi_b(\mathbf{r})), \quad (2.99)$$

where ϕ_a is 1s spatial orbital on atom A, ϕ_b is 1s spatial orbital on atom B and S is the overlap integral given as $\langle \phi_a | \phi_b \rangle$.

At bond dissociation, $S \rightarrow 0$, so the gerade orbital becomes

$$\psi_{\sigma_g}(\mathbf{r}) = \frac{1}{\sqrt{2}}(\phi_a(\mathbf{r}) + \phi_b(\mathbf{r})). \quad (2.100)$$

In RHF, both electrons occupy the same spatial orbital and take opposite spins. So, the RHF wave function will be

$$|\Phi_0^{RHF}\rangle = |\sigma_g \bar{\sigma}_g\rangle = \psi_{\sigma_g}(\mathbf{1}) \psi_{\sigma_g}(\mathbf{2}) \frac{\alpha(1)\beta(2) - \beta(1)\alpha(2)}{\sqrt{2}}. \quad (2.101)$$

Then, taking the spatial part only as this part affects the energy,

$$\begin{aligned} \psi_{\sigma_g}(\mathbf{1}) \psi_{\sigma_g}(\mathbf{2}) &= \frac{1}{2}(\phi_a(1) + \phi_b(1))(\phi_a(2) + \phi_b(2)) \\ &= \frac{1}{2}[\phi_a(1)\phi_a(2) + \phi_a(1)\phi_b(2) + \phi_b(1)\phi_a(2) + \phi_b(1)\phi_b(2)]. \end{aligned} \quad (2.102)$$

Representing in occupation kets,

$$\begin{aligned} |\Phi_0^{RHF}\rangle &= \frac{1}{2}(|a\bar{a}\rangle + |a\bar{b}\rangle + |b\bar{a}\rangle + |b\bar{b}\rangle) \text{ upto spin singlet factor.} \\ &= \frac{1}{2}(|a\bar{b}\rangle + |b\bar{a}\rangle + |a\bar{a}\rangle + |b\bar{b}\rangle). \end{aligned} \quad (2.103)$$

In the equation 2.103, the first two terms form covalent part and the last two terms for the ionic part and they are expressed as,

$$|cov\rangle = \frac{1}{2}(|a\bar{b}\rangle + |b\bar{a}\rangle), \quad (2.104)$$

$$|ion\rangle = \frac{1}{2}(|a\bar{a}\rangle + |b\bar{b}\rangle). \quad (2.105)$$

While ($R \rightarrow \infty$), in the covalent configuration, there is no electron-electron repulsion between the electrons, thus the energy $E_{cov} = h_{aa} + h_{bb}$. By symmetry, $E_{cov} = 2h_{aa}$.

Similarly, as we examine the ionic configuration, we know that in RHF, both electrons sit on the same atom (spatial orbital). So, there is a Coulomb repulsion on the same atom and given as J_{aa} and one-electron part given as $2h_{aa}$ as for the case of covalent configuration. It gives $E_{ion} = 2h_{aa} + J_{aa}$ for $|a\bar{a}\rangle$. We can obtain the similar form for $|b\bar{b}\rangle$ too, yet only one would be valid for consideration.

Thus, the RHF energy as expectation value of the Hamiltonian with respect to $|\Phi_0^{RHF}\rangle$ is given as

$$\begin{aligned} E_{RHF} &= \langle \Phi_0^{RHF} | \hat{H} | \Phi_0^{RHF} \rangle \approx \frac{1}{2}E_{cov} + \frac{1}{2}E_{ion}, \\ &= \frac{1}{2}(2h_{aa}) + \frac{1}{2}(2h_{aa} + J_{aa}), \\ &= 2h_{aa} + \frac{1}{2}J_{aa}. \end{aligned} \quad (2.106)$$

The bond dissociation of H_2 is expected to yield two independent H atoms where the total energy of the system would be the sum of one electron-term energy contribution h_{aa} . However RHF puts both electrons on one atom, generating extra term $\frac{1}{2}J_{aa}$. This term is the spurious term in RHF energy and inaccurately increases the energy of the system.

However, in the case of UHF, the electrons are allowed to occupy different spatial orbitals, $\phi^\alpha(\mathbf{r}) \neq \phi^\beta(\mathbf{r})$. This allows the Hartree–Fock equations to spontaneously localize the electrons on separate atoms when the bond is stretched. At bond dissociation,

$$|\Phi_0^{UHF}\rangle = |a\bar{b}\rangle \text{ or } |b\bar{a}\rangle. \quad (2.107)$$

These solutions are pure covalent structures and contain no ionic terms as before. Also, $J_{ab} \rightarrow 0$, $K_{ab} \rightarrow 0$, and $E_{UHF} = 2h_{aa}$ or $2h_{bb}$. Hence, UHF is the proper method for open-shell systems including bond dissociation.

2.5 Quantum simulation

Quantum simulation is the most generally understood as simulating the Hamiltonian dynamics of a quantum system which is beyond classical capabilities, over time according to the Schrödinger equation 2.34. Overall it involves the problems like finding the ground state, excited states, real and imaginary time dynamics, thermal properties, correlation functions, quantum phase transitions, etc. for systems including molecules, lattice models, quantum field theories and open quantum systems. Quantum simulation covers simulating the quantum system while mimicing its behavior by quantum control called analog simulation. However, throughout the thesis we use the term, ‘Quantum simulation’ to refer digital simulation which uses a sequence of unitary gates in quantum algorithms to simulate the system on a universal quantum computer.

Since many physical system of interest are described by a Hamiltonian, Hamiltonian simulation is the central computational primitive of quantum simulation. Solving the Hamiltonian to get the dynamics using quantum computer provides exponential speedups for chemical simulations and quantum field theories. Not just that, the techniques used can be important for solving linear systems, semi-definite programming, ordinary differential equations and possibly more.

The main objective of Hamiltonian simulation is to create a quantum circuit that efficiently approximates the behavior of the unitary time evolution operator

$$U(t) = e^{-i\hat{H}t} = 1 - i\hat{H}t - \frac{\hat{H}^2 t^2}{2!} + \dots \quad (2.108)$$

The matrix \hat{H} corresponding to a system of n -qubits is of the dimension $2^n \times 2^n$, which is exponentially large and becomes very hard to simulate for classical computers. Even when we figure out some ways to efficiently compute \hat{H} using known structures of the Hamiltonian, getting the time evolution is still a BQP-complete problem. With Hamiltonian simulation on a quantum computer, we aim to do the time evolution of a physical system using unitary evolution with a gate sequence of polynomial size on a quantum computer such that the final state on the computer is logically equivalent to the final state we get in the physical system up to error ϵ for any value of ϵ and t specified.

Let us have a fundamental understanding of how we do quantum simulation on a gate based quantum computer for a very basic Hamiltonian given by $\hat{H} = \alpha X \otimes X$. The given Hamiltonian matrix looks like the following

$$\alpha \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}. \quad (2.109)$$

The steps to simulate this Hamiltonian include the following:

1. Diagonalize \hat{H} as, $\hat{H} = \alpha(H \otimes H)(Z \otimes Z)(H \otimes H)$.

2. Compute the eigen values of the diagonalized matrix as, $(Z \otimes Z) |ij\rangle = (-1)^{i+j} |ij\rangle$.
3. Create a quantum circuit to implement the unitary U that does $|ij\rangle \rightarrow e^{-i\alpha(-1)^{i+j} Zt} |ij\rangle$.
4. Translate it back to the computational basis by applying H gates.

The complete circuit for simulating $\hat{H} = \alpha X \otimes X$ is shown in figure 2.20.

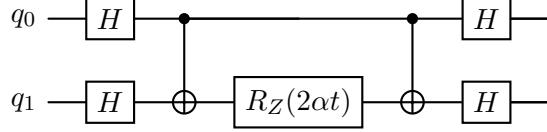


Figure 2.20: Circuit for simulating the Hamiltonian given as $\alpha X \otimes X$.

2.5.1 Suzuki–Trotter simulation

Suzuki–Trotter simulation [31] is the first explicit algorithm for simulating the Hamiltonian dynamics and still remains a good choice for various problem instances. The idea behind this approach is breaking the larger unitary problem into a sequence of easy problems that are easy to implement.

\hat{H} can be represented as the sum of m Hamiltonian terms as

$$\hat{H} = \sum_{j=1}^m a_j H_j. \quad (2.110)$$

The time evolution operator corresponding to the Hamiltonian in equation (2.110) is given as

$$U(t) = e^{-i \sum_{j=1}^m H_j t}. \quad (2.111)$$

and

$$\begin{aligned} e^{-i\hat{H}t} &= e^{-i(H_1+H_2+\dots+H_m)t} \\ &= e^{-iH_1t}e^{-iH_2t}\dots e^{-iH_mt} \text{ for all } t \text{ if } [H_j, H_k] = 0 \text{ for all } j, k. \end{aligned} \quad (2.112)$$

But if $[H_j, H_k] \neq 0$, $e^{-i(H_1+H_2+\dots+H_m)t} \neq e^{-iH_1t}e^{-iH_2t}\dots e^{-iH_mt}$, and it follows Baker–Campbell–Hausdorff (BCH) formula which states

$$e^{(A+B)\Delta t} = e^{A\Delta t}e^{B\Delta t}e^{-\frac{1}{2}[A,B]\Delta t^2} + O(\Delta t^3) \quad (2.113)$$

for small time steps Δt .

The first order Trotter formula is given by

$$e^{-i\hat{H}t} = e^{-i \sum_{j=1}^m H_j t} = \left(\prod_{j=1}^m e^{-iH_j \frac{t}{r}} \right)^r + O\left(\frac{t^2}{r}\right), \quad (2.114)$$

where r is the number of Trotter steps and the approximation can be made better with higher values of r . The first order trotter formula approximates the time evolution with an error scaling of the order $O\left(\frac{t^2}{r}\right)$, depending on the number of Trotter steps and the total time of evolution.

The second order Trotter formula is also called Suzuki–Trotter formula and it provides the error scaling of $O\left(\frac{t^3}{r^2}\right)$, and is given as

$$\begin{aligned} e^{-i\hat{H}t} &= \left(\prod_{j=1}^m e^{-iH_j \frac{t}{2r}} \prod_{j=m}^1 e^{-iH_j \frac{t}{2r}} \right)^r + O(r\Delta t^3). \\ \Rightarrow \left\| e^{-i\hat{H}t} - \left(\prod_{j=1}^m e^{-iH_j \frac{t}{2r}} \prod_{j=m}^1 e^{-iH_j \frac{t}{2r}} \right)^r \right\| &= O\left(r\left(\frac{t}{r}\right)^3\right) = O\left(\frac{t^3}{r^2}\right). \end{aligned} \quad (2.115)$$

The second order formula provides square root improvement as $r = O\left(\frac{t^{\frac{3}{2}}}{\epsilon^{\frac{1}{2}}}\right)$ compared to the first order decomposition $r = O\left(\frac{t^2}{\epsilon}\right)$.

For p-th order formula, the number of trotter steps required scales as $r = O\left(\frac{t^{1+\frac{1}{p}}}{\epsilon^{\frac{1}{p}}}\right)$.

For the first order, the total number of unitaries needed to simulate H is given as $O\left(\frac{\Lambda^2 M^3 t^2}{\epsilon}\right)$ where $\Lambda = \max_k h_k$ and M is the number of terms in the Hamiltonian.

The cost with Suzuki–Trotter simulations scales with commutators of terms in the Hamiltonian. It also works with time dependent Hamiltonian and needs no ancilla for simulations. However, the scaling is not optimal and the circuit depth increases very rapidly.

2.5.2 qDrift

qDrift is quantum stochastic drift protocol, which is a randomized compiler for Hamiltonian simulation where the gate probabilities are proportional to the strength of a corresponding term in the Hamiltonian [32]. For the Hamiltonian \hat{H} expressed as

$$\hat{H} = \sum_{k=1}^M h_k H_k, \quad (2.116)$$

with $h_k > 0$, we have $\lambda = \sum_{k=1}^M h_k$ as the upper bound for the largest singular value of \hat{H} and $\|H_k\| \leq 1$.

Using the simplest Trotter formula dividing $U = e^{-i\hat{H}t}$ into M segments, a single unitary $U_k = e^{-i\hat{H}_k \Delta t}$.

Then, we take an operator V_k which approaches U_k in the large limit of k . It is expressed as

$$V_k = \prod_{j=1}^L e^{-i\tau H_{k_j}}; \quad k = \{k_1, k_2, \dots, k_L\}, \quad (2.117)$$

where k_j takes 1, ..., M randomly.

Here, we will just evaluate the first order approximation by qDrift and leave the rest for the reader to check [32].

We want to evaluate the expectation of the trajectories taken by each path in qDrift. The expectation value of random operator V_k , $E[V_k]$ is obtained with the theory of ensemble of mixed states. Assuming $|\phi_k\rangle$ appears with probability P_k , we measure the system as,

$$\begin{aligned} \text{Prob}(\#k) &= \sum_i \text{Prob} (\#k \mid \text{state} = |\phi_i\rangle) \cdot \text{Prob}(|\phi_i\rangle) \\ &= \sum_i \langle \phi_i | P_k | \phi_i \rangle \cdot P_i \\ &= \text{Tr}[\sum_i P_i |\phi_i\rangle \langle \phi_i| \cdot P_k] \\ &= \text{Tr}(\rho P_k), \end{aligned} \quad (2.118)$$

where $\rho = \sum_i P_i |\phi_i\rangle \langle \phi_i|$ is the state describing the ensemble.

Assuming all $|\phi_i\rangle$ evolve via the same \hat{H} such that $|\dot{\phi}_i\rangle = -i\hat{H}|\phi_i\rangle \forall i$, the evolved state $\dot{\rho}$ is given as

$$\begin{aligned} \dot{\rho} &= \frac{d}{dt} \left(\sum_i P_i |\phi_i\rangle \langle \phi_i| \right) \\ &= \sum_i P_i \left(|\dot{\phi}_i\rangle \langle \phi_i| + |\phi_i\rangle \langle \dot{\phi}_i| \right) \\ &= \sum_i P_i (-iH |\phi_i\rangle \langle \phi_i| + i |\phi_i\rangle \langle \phi_i| H) \\ &= -iH \left(\sum_i P_i |\phi_i\rangle \langle \phi_i| \right) + i \left(\sum_i P_i |\phi_i\rangle \langle \phi_i| \right) H \\ &= -iH\rho + i\rho H = -i[H, \rho]. \end{aligned} \quad (2.119)$$

Thus, $\dot{\rho} = -i[H, \rho]$ and it gives $\rho(t) = e^{-iHt}\rho(0)e^{iHt}$.

Now lets find the expectation value of each path taken by qDrift as

$$\begin{aligned}
E[e^{-i\tau H} \rho e^{i\tau H}] &= \sum_k P_k e^{-i\tau H_k} \rho e^{i\tau H_k}; \quad P_k = \frac{h_k}{\lambda} = \frac{h_k}{\sum_k h_k} \\
&= \sum_k \frac{h_k}{\lambda} \left(1 - i\tau H_k - \frac{\tau^2 H_k^2}{2}\right) \rho \left(1 + i\tau H_k - \frac{\tau^2 H_k^2}{2}\right) + O(\tau^3) \\
&= \sum_k \frac{h_k}{\lambda} (\rho + i\tau \rho H_k - i\tau H_k \rho) + O(\tau^2) \\
&= \rho - \frac{i\tau}{\lambda} \sum_k h_k [H_k, \rho] + O(\tau^2) \\
&= \rho - \frac{i\tau}{\lambda} [H, \rho] + O(\tau^2).
\end{aligned} \tag{2.120}$$

The target trajectory in time t/L is given by

$$\begin{aligned}
e^{-iH\frac{t}{L}} \rho e^{iH\frac{t}{L}} &= \left(1 - \frac{itH}{L}\right) \rho \left(1 + \frac{itH}{L}\right) + O\left(\frac{t^2}{L^2}\right) \\
&= \rho - i\frac{t}{L} [H, \rho] + O\left(\frac{t^2}{L^2}\right).
\end{aligned} \tag{2.121}$$

Comparing equations 2.120 and 2.121,

$$\frac{t}{L} = \frac{\tau}{\lambda} \Rightarrow \tau = \frac{t\lambda}{L}. \tag{2.122}$$

Thus, the first order approximation using qDrift is similar as the target state when $\tau = \frac{t\lambda}{L}$.

Evaluating the diamond norm¹ of exact \mathcal{U}_ρ and the qDrift state \mathcal{E}_ρ gives

$$\begin{aligned}
\|\mathcal{U} - \mathcal{E}\|_\diamond &= \left\| \sum_{n=2}^{\infty} \frac{(t/L)^n}{n!} \mathcal{L}^n - \sum_n \frac{h_k}{\lambda} \sum_{n=2}^{\infty} \frac{\tau^n \mathcal{L}_k^n}{n!} \right\|_\diamond \\
&\leq \sum_{n=2}^{\infty} \frac{(t/L)^n}{n!} (2\lambda)^n + \sum_k \frac{h_k}{\lambda} \sum_{n=2}^{\infty} \frac{(2\tau)^n}{n!}; \quad [\because \|\mathcal{L}\|_\diamond^n \leq (2\lambda)^n, \quad \|\mathcal{L}_k\|_\diamond^n \leq 2^n] \\
&= 2 \sum_{n=2}^{\infty} \frac{(2\tau)^n}{n!}; \quad \left[\because \left(\frac{2t\lambda}{L}\right)^n = (2\tau)^n, \sum_k \frac{h_k}{\lambda} = 1 \right].
\end{aligned} \tag{2.123}$$

Here, \mathcal{L} is the Liouvillian superoperator and it maps density matrices to density matrices. For a pure unitary channel, it takes the form of commutator as $\mathcal{L}(\rho) = i(H\rho - \rho H)$. From equation 2.123, we get the upper bound as $(2\tau)^2 = 4(\frac{t\lambda}{L})^2 = \frac{4t^2\lambda^2}{L^2}$. This gives the number of unitaries required as $O\left(\frac{t^2\lambda^2}{\epsilon}\right)$. This removes the dependence of the implementation cost on the number of terms M in the Hamiltonian on contrary to the first order Trotter method.

¹The diamond norm, $\|\Phi - \Psi\|_\diamond$ measures the worst case measure of the distinguishability between two quantum channels Φ and Ψ . It is defined as the maximum trace distance between the outputs of two channels, which are optimized over all input states and all possible entanglement with an ancillary system.

2.5.3 Other methods

While trotter methods like Suzuki-Trotter and qDrift do not need any ancilla for simulating the Hamiltonian, there are other methods that use ancilla, and provide better asymptotic performance including Linear Combination of Unitaries (LCU) [33], Qubitization [34], Tensor Hypercontraction [35], Spectrum Amplification [36] and very recent Sum-of-squares Spectral Amplification (SOSSA) [37]. During the time of writing this thesis, the state-of-the-art algorithm for Hamiltonian simulation is SOSSA, needing 1137(1459) qubits with 3.41×10^8 (9.99×10^8) toffoli gates to simulate FeMoco-54 [38](Femoco-76 [39]) within hours. Since we do not use or deal with these algorithms within the scope of this thesis, we leave further readings for the reader from the sources given if necessary.

2.6 Quantum imaginary time evolution

2.6.1 Imaginary time evolution

Imaginary time is a mathematical concept used in statistical mechanics, quantum field theory, and cosmology. It involves rotating the time axis in complex space, from real time t to imaginary time τ as $t \rightarrow i\tau$. Exponential integrals such as $e^{i\hat{H}t}$ oscillate, making it often difficult to evaluate their long-term properties. Rotating to imaginary time maps the oscillations to a smooth function with decay rate dependent on the size of the integral components. Thus, performing ITE allows the function to decay towards the lowest energy eigen state of the system evolving in time.

The imaginary-time Schrödinger equation takes the form

$$-\partial_\tau |\Phi(\tau)\rangle = \hat{H}|\Phi(\tau)\rangle. \quad (2.124)$$

The ground state $|\Psi_0\rangle$ of the Hamiltonian \hat{H} , with $\langle\Phi(0)|\Psi_0\rangle \neq 0$, is given by the long-term limit of (2.124),

$$|\Psi_0\rangle = \lim_{\tau \rightarrow \infty} \frac{|\Phi(\tau)\rangle}{\| |\Phi(\tau)\rangle \|}. \quad (2.125)$$

2.6.2 Unitary approximation

QITE decomposes the Hamiltonian into $\hat{H} = \sum_m \hat{h}[m]$, where all $\hat{h}[m]$ are at most j -local, and Trotterizes the generated evolution,

$$e^{-\tau\hat{H}} = (e^{-\Delta\tau\hat{h}[1]} e^{-\Delta\tau\hat{h}[2]} \dots)^{\tau/\Delta\tau} + \mathcal{O}(\Delta\tau). \quad (2.126)$$

If we consider j to be non-local, but global in the Hamiltonian, we can represent the imaginary time propagator for time $\Delta\tau$ as $e^{-\Delta\tau\hat{H}}$, which is a non-unitary operator.

If $|\Psi\rangle$ is the initial state, upon an imaginary time step, the state becomes

$$|\Psi\rangle \rightarrow e^{-\Delta\tau\hat{H}}|\Psi\rangle = |\Psi'\rangle. \quad (2.127)$$

We approximate the normalized non-unitary by a unitary operator with real time evolution as,

$$|\Psi\rangle \rightarrow e^{-i\Delta\tau\hat{A}}|\Psi\rangle = |\Psi''\rangle, \quad (2.128)$$

where \hat{A} is a Hermitian operator acting on all system qubits, and $\hat{A} = \sum_j a_j \sigma_j$; $j \in \{0, 1, 2, 3\}$, $\sigma \in \{I, X, Y, Z\}$, $a_j \in \mathbb{R}$ and the coefficients a_j are obtained from solving the linear system in equation 2.138.

The norm of $|\Psi'\rangle$ in equation (2.127) shrinks, thus it has to be normalized as,

$$|\bar{\Psi}'\rangle = \frac{e^{-\Delta\tau\hat{H}}|\Psi\rangle}{\|e^{-\Delta\tau\hat{H}}|\Psi\rangle\|}. \quad (2.129)$$

For convenience, we write $c^{\frac{1}{2}} = \|e^{-\Delta\tau\hat{H}}|\Psi\rangle\| = \sqrt{\langle\Psi|e^{-2\Delta\tau\hat{H}}|\Psi\rangle} \rightarrow c = \langle\Psi|e^{-2\Delta\tau\hat{H}}|\Psi\rangle$.

If we define $|\Delta_0\rangle = \frac{|\bar{\Psi}'\rangle - |\Psi\rangle}{\Delta\tau}$ and $|\Delta\rangle = \frac{|\Psi''\rangle - |\Psi\rangle}{\Delta\tau}$, we want to minimize $\|\Delta_0 - \Delta\|^2$. Let us define the function to minimize for real a as,

$$f = \langle(\Delta_0 - \Delta)|(\Delta_0 - \Delta)\rangle \quad (2.130)$$

$$= \langle\Delta_0|\Delta_0\rangle - \langle\Delta_0|\Delta\rangle - \langle\Delta|\Delta_0\rangle + \langle\Delta|\Delta\rangle. \quad (2.131)$$

Constructing preliminaries for simplifying equation (2.131),

First order expansion of exponential terms in $|\Delta\rangle$ gives,

$$|\Delta\rangle = \frac{e^{-i\Delta\tau\hat{A}}|\Psi\rangle - |\Psi\rangle}{\Delta\tau} = \frac{(|\Psi\rangle - i\Delta\tau\hat{A}|\Psi\rangle) - |\Psi\rangle}{\Delta\tau} = -i\hat{A}|\Psi\rangle = -i\sum_I a_I \sigma_I |\Psi\rangle. \quad (2.132)$$

Using the simplified expression in equation (2.132), we obtain the terms as,

$$\langle\Delta|\Delta\rangle = i\sum_I a_I^* \sigma_I^\dagger \langle\Psi|(-i)\sum_J a_J \sigma_J |\Psi\rangle = \sum_{IJ} a_I^* a_J \langle\Psi|\sigma_I^\dagger \sigma_J |\Psi\rangle \quad (2.133)$$

$$\langle\Delta_0|\Delta\rangle = \langle\Delta_0|(-i\sum_I a_I \sigma_I)|\Psi\rangle = -i\sum_I a_I \langle\Delta_0|\sigma_I|\Psi\rangle \quad (2.134)$$

$$\langle\Delta|\Delta_0\rangle = i\sum_I a_I^* \langle\Psi|\sigma_I^\dagger|\Delta_0\rangle \quad (2.135)$$

Thus, we obtain the total minimizing function as,

$$\begin{aligned} f &= f_0 - \left(-i\sum_I a_I \langle\Delta_0|\sigma_I|\Psi\rangle \right) - \left(i\sum_I a_I^* \langle\Psi|\sigma_I^\dagger|\Delta_0\rangle \right) + \left(\sum_{IJ} a_I^* a_J \langle\Psi|\sigma_I^\dagger \sigma_J |\Psi\rangle \right) \\ &= f_0 + i\sum_I a_I \langle\Delta_0|\sigma_I|\Psi\rangle - i\sum_I a_I^* \langle\Psi|\sigma_I^\dagger|\Delta_0\rangle + \sum_{IJ} a_I^* a_J \langle\Psi|\sigma_I^\dagger \sigma_J |\Psi\rangle \\ &= f_0 + \sum_I b_I a_I + \sum_{IJ} a_I^* S_{IJ} a_J \end{aligned} \quad (2.136)$$

where, $f_0 = \langle\Delta_0|\Delta_0\rangle$, $S_{IJ} = \langle\Psi|\sigma_I^\dagger \sigma_J |\Psi\rangle$ and, $b_I = i\langle\Delta_0|\sigma_I|\Psi\rangle - i\langle\Psi|\sigma_I^\dagger|\Delta_0\rangle$.

Minimizing equation (2.136) for real a , we obtain

$$\frac{\partial f}{\partial a_I} = b_I + \sum_J (S_{IJ} + S_{JI})a_J = 0. \quad (2.137)$$

This gives

$$(S + S^\dagger)a = -b. \quad (2.138)$$

Finally, we solve the linear form in equation (2.138) to find the real entries of a .

Chapter 3

Design rationale and implementation

Parts of this chapter are reproduced with permission from Poudel *et al.* [12].

3.1 Motivation

3.1.1 Problems with simulating strongly-correlated systems

The conventional practice for GSEE in the noisy intermediate-scale quantum (NISQ) era [40] is to use VQE with the Unitary Coupled Cluster (UCC) ansatz with the spin-restricted HF (RHF) wave function as the reference [41]. However, RHF is a poor approximation for open shell systems with antiferromagnetic couplings between unpaired electrons. Consequently, consideration of higher-order excitation operators is necessary to obtain an accurate ground-state energy in VQE-UCC. In the real- and imaginary-time evolution-based ASP and QITE, the convergence speed is strongly controlled by the energy gap between the ground and the excited states. Narrowing energy gap results in decreasing convergence rate in the GSEE of open shell systems.

Several approaches have been proposed to tackle this problem. Instead of single configurational HF wave functions, techniques using multi-configurational wave functions such as Complete Active Space Self-Consistent Field (CASSCF) with smaller active space have been well studied. For example, multireference UCC methods like Multireference-Unitary Coupled Cluster with partially Generalized Singles and Doubles (MR-UCCpGSD) ansatz [42] were proposed for VQE. An approach using the CASSCF wave function as the initial wave function has also been investigated in ASP [43]. These methods greatly improve the convergence, but solving CASSCF itself becomes intractable when a large active space is required to capture strong correlations. The need for deep quantum circuits to encode multiconfigurational reference wave functions is another problem in quantum computation. It is highly desirable to develop theoretical methods that are applicable to strongly correlated systems, with shallow quantum circuits for initial state preparation.

3.2 The broken-symmetry wave function

The electronic configuration of the highest spin state with spin quantum number $S = N/2$ with spin magnetic quantum number $|M| = S$ can be described by a single determinant. On the other hand, low-spin ($S < N/2$) states and $|M| \neq S$ sublevels of the highest spin state show inherent multi-determinant character. This prevents the application of conventional single-determinant methods such as Density Functional Theory (DFT) to compute open shell low-spin states. BS wave function addresses this issue by using a linear combination of the high- and low-spin wave functions to generate a spin-mixed single configurational wave function. BS wave functions have been widely used in spin chemistry to calculate exchange interactions between unpaired electrons that determine the energy gap between spin states. On the quantum computing side, the BS wave function is described by a single electron configuration, which has the advantage that state preparation can be done with Pauli-X gates only, and the cost of preparing the BS wave function is equivalent to that of the RHF wave function.

The branching diagram shown in Fig. 3.1 is a blueprint for constructing spin eigenfunctions [44]. It consists of circles on the vertices and diagonal upward (45°) and downward (-45°) lines connecting vertices. The upward (downward) line represents adding an electron spin to increase (decrease) spin quantum number. The number in the circle represents the number of possible pathways to arrive the vertex and it corresponds to the number of spin eigenfunctions for given N and S . The BS wave function of an N -spin system with magnetic quantum number M can be written as the linear combinations of the spin eigenfunctions as follows:

$$|\text{BS}(N, M)\rangle = \sum_{j=0}^{N/2} a_j |N, S = j, M\rangle, \quad (3.1)$$

where the probability amplitudes a_j can be obtained from the branching diagram shown in figure 3.1.

Here, $g_{N,j}$ is the number of spin eigenfunctions (values in the circle in the branching diagram) of N -spin system with $S = j$, and $G_N = \sum_k g_{N,k}$. Using this equation, we can easily derive the expanded form of the BS wave function in terms of spin eigenfunctions. For example, for two-spin system, the BS wave function $|\alpha\beta\rangle$ can be written as the linear combination of $|2, 1, 0\rangle$ and $|2, 0, 0\rangle$ states, and the coefficients are $\frac{1}{\sqrt{2}}$ for both. Therefore,

$$|\alpha\beta\rangle = \frac{1}{\sqrt{2}}|2, 0, 0\rangle + \frac{1}{\sqrt{2}}|2, 1, 0\rangle. \quad (3.2)$$

A six-spin BS wave function of $M = 0$ can be derived as:

$$\begin{aligned} |\alpha\alpha\alpha\beta\beta\beta\rangle &= \frac{1}{\sqrt{20}}|6, 3, 0\rangle + \frac{1}{2}|6, 2, 0\rangle + \\ &\quad \frac{3}{\sqrt{20}}|6, 1, 0\rangle + \frac{1}{2}|6, 0, 0\rangle. \end{aligned} \quad (3.3)$$

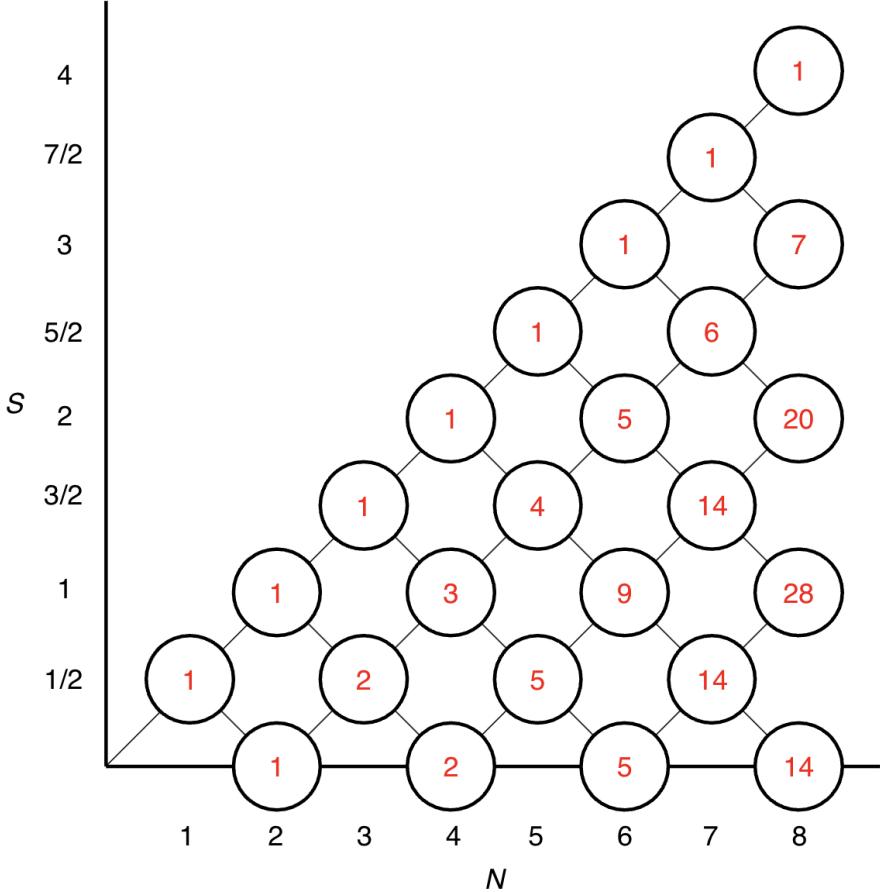


Figure 3.1: Branching diagram for the construction of spin eigenfunctions.

Let us consider two spin eigenfunctions $|2, 0, 0\rangle$ and $|2, 1, 0\rangle$. Taking their equal superposition generates the spin state $|\alpha\beta\rangle$ as follows,

$$|\alpha\beta\rangle = \frac{1}{\sqrt{2}}|2, 0, 0\rangle + \frac{1}{\sqrt{2}}|2, 1, 0\rangle. \quad (3.4)$$

This wave function corresponds to a BS state of the H_2 molecule. The notation $|\alpha\beta\rangle$ in (3.4) does not represent spin-up and spin-down state of two electrons. Rather, it denotes the occupancy of active molecular orbitals. The BS state in (3.4) describes a state where the lower orbital is occupied by a single spin-up electron and the second orbital contains a single spin-down electron. More generally, a molecule with k active orbitals is denoted by $|c_1 c_2 \dots c_k\rangle$, where $c_i \in \{2, \alpha, \beta, 0\}$. Occupancy $c_i = 2$ means the orbital contains paired electrons, while $c_i = 0$ is an unoccupied orbital.

Let us consider a N_2 molecule which has triple bonds (two π bonds and a σ bond) between two N atoms. We represent the BS wave function corresponding to two π -bond dissociation as BS2 and both π and σ bond dissociation as BS3. Fig. 3.2 compares the

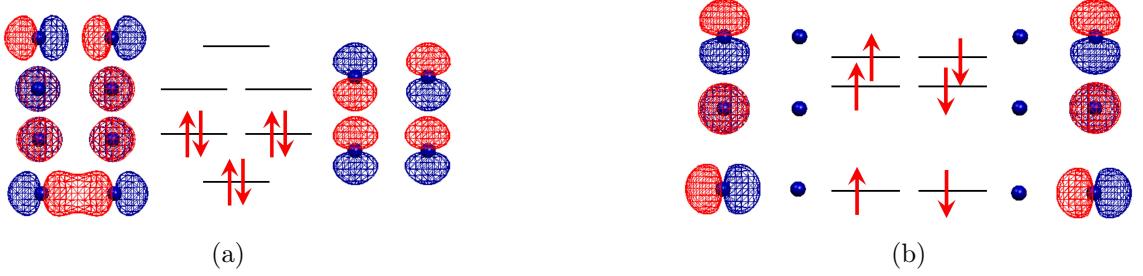


Figure 3.2: Active orbitals of the N₂ molecule. The blue dots represent N atoms. (a) RHF canonical orbitals. Arrows specify the RHF electronic configuration. (b) Localized orbitals constructed from the natural orbitals computed at the BS-UHF level. Arrows specify the electron occupancies of the BS3. (Figures were generated using Molekel4.3 [45].)

electronic configuration of RHF and BS3 active orbitals. In the active space consisting of valence σ and π orbitals, the electron configuration of the BS2 and BS3 are represented as $|2\alpha\alpha\beta\beta0\rangle$ and $|\alpha\alpha\alpha\beta\beta\beta\rangle$, respectively. BS2 and BS3 wave functions are described as the sum of states of different spin multiplicities with zero spin magnetic quantum number,

$$|2\alpha\alpha\beta\beta0\rangle = \frac{1}{\sqrt{6}}|4, 2, 0\rangle + \frac{1}{\sqrt{2}}|4, 1, 0\rangle + \frac{1}{\sqrt{3}}|4, 0, 0\rangle, \quad (3.5)$$

$$\begin{aligned} |\alpha\alpha\alpha\beta\beta\beta\rangle = & \frac{1}{\sqrt{20}}|6, 3, 0\rangle + \frac{1}{2}|6, 2, 0\rangle + \frac{3}{\sqrt{20}}|6, 1, 0\rangle \\ & + \frac{1}{2}|6, 0, 0\rangle. \end{aligned} \quad (3.6)$$

Again, the probability amplitudes are obtained from the branching diagram in figure 3.1.

3.3 Spin penalty to Hamiltonian

The second quantized fermionic electronic Hamiltonian is represented as

$$\hat{H} = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s + E_{nuc}, \quad (3.7)$$

where h_{pq} and h_{pqrs} are one- and two-electron integrals respectively, and E_{nuc} is the constant nuclear repulsion term.

For each spatial orbital $i = 0, \dots, n_{mo} - 1$, we define the spin-orbital indices as $i\alpha = 2i$ and $i\beta = 2i + 1$.

Then the number operators are defined as, $\hat{n}_{i\alpha} = a_{i\alpha}^\dagger a_{i\alpha}$ and $\hat{n}_{i\beta} = a_{i\beta}^\dagger a_{i\beta}$. In general, a_p^\dagger and a_p are creation and annihilation operators that create and annihilate an electron, respectively, in a spin orbital p .

Similarly, while the local spin-operators are defined as

$$\hat{S}_z(i) = \frac{1}{2}(\hat{n}_{i\alpha} - \hat{n}_{i\beta}), \quad \hat{S}_+(i) = a_{i\alpha}^\dagger a_{i\beta}, \quad \hat{S}_-(i) = a_{i\beta}^\dagger a_{i\alpha}, \quad (3.8)$$

we represent operator $\hat{\mathbf{S}}^2$ in a beautiful form as

$$\hat{\mathbf{S}}^2 = \frac{1}{2} (\hat{S}_+ \hat{S}_- + \hat{S}_- \hat{S}_+) + \hat{S}_z^2, \quad (3.9)$$

$$= \sum \left[\frac{1}{2} (\hat{S}_+(i) \hat{S}_-(j) + \hat{S}_-(i) \hat{S}_+(j)) + \hat{S}_z(i) \hat{S}_z(j) \right]. \quad (3.10)$$

Expanding equation (3.10) in terms of creation and annihilation operators,

$$\begin{aligned} \hat{\mathbf{S}}^2 = & \frac{1}{4} \sum_{i,j} \left(a_{i\alpha}^\dagger a_{i\alpha} a_{j\alpha}^\dagger a_{j\alpha} - a_{i\alpha}^\dagger a_{i\alpha} a_{j\beta}^\dagger a_{j\beta} \right) \\ & + \frac{1}{4} \sum_{i,j} \left(-a_{i\beta}^\dagger a_{i\beta} a_{j\alpha}^\dagger a_{j\alpha} + a_{i\beta}^\dagger a_{i\beta} a_{j\beta}^\dagger a_{j\beta} \right) \\ & + \frac{1}{2} \sum_{i,j} \left(a_{i\alpha}^\dagger a_{i\beta} a_{j\beta}^\dagger a_{j\alpha} + a_{i\beta}^\dagger a_{i\alpha} a_{j\alpha}^\dagger a_{j\beta} \right). \end{aligned} \quad (3.11)$$

In this work, we shift our symmetry-broken Hamiltonian by some scaling component introduced by the eigen value of $\hat{\mathbf{S}}^2$ operator in equation (3.11). Thus, our working Hamiltonian \hat{H}' will be the symmetry-broken Hamiltonian with penalty,

$$\hat{H}' = \hat{H} + c \hat{\mathbf{S}}^2, \quad (3.12)$$

where c is the strength of the penalty. We take $c = 1$ for most of our simulations. This is the noble work in this thesis.

3.3.1 Intuition behind symmetry breaking and penalty

The addition of the penalty term adds an energy contribution proportional to $cS(S + 1)$ for each spin component present in the wave function. Low spins suffer a small penalty in comparison to large spins, leading to an increase in the energy gap between the singlet ground state and the excited states belonging to different spin multiplicities. Ultimately, this leads to faster convergence to the ground state.

It is important to note that the HF wave function is the eigenfunction of the $\hat{\mathbf{S}}^2$ operator, and in the bond dissociation limit of the H₂ molecule, the HF wave function is expressed as a superposition of the singlet ground state and the singlet excited state. Since the excited state of the same spin multiplicity as the ground state contributes to the wave function, it is difficult to apply a penalty term to lift the quasi-degeneracy.

In contrast, the BS wave function for H₂ molecule in the bond dissociation limit is described as the linear combination of the singlet ground state and the triplet excited state

wave functions. In this case, introduction of the $\hat{\mathbf{S}}^2$ operator to the Hamiltonian raises the energy of the triplet state lifting the quasi-degeneracy of the two states. We emphasize that the introduction of the $\hat{\mathbf{S}}^2$ penalty term is essential for QITE with the BS wave function. Without the penalty term, BS-QITE usually converges slower than RHF-QITE, as the S_0-T_1 energy gap is generally smaller than the S_0-S_1 energy difference, where S_0 , S_1 , and T_1 represent the lowest singlet state, the first excited singlet state, and the lowest triplet state, respectively.



(a) Singlet ground state and singlet first excited state energies with HF wave function.
 (b) Singlet ground state and triplet first excited state energies with BS wave function and penalty.

Figure 3.3: Representational comparison of energy gap between the ground state and the first excited state corresponding to delocalized HF and penalty added localized BS wave function.

Fig. 3.3a shows the schematic view of the energy gap ($E_1 - E_0$) between the singlet ground state and singlet excited state in the H_2 molecule. In the vicinity of the equilibrium geometry, the energy gap is large, but the gap becomes smaller when the covalent bond dissociation takes place. In the bond dissociation limit, the ground-state wave function is well approximated by the linear combination of the RHF configuration and two-electron excited configuration from the RHF state, $|\Psi_0\rangle = \frac{1}{\sqrt{2}}(|20\rangle - |02\rangle)$. In contrast, when the molecular orbitals localized onto spin sites are used as the wave function expansion, the ground state wave function is approximated by a linear combination of open shell configurations: $|\Psi_0\rangle = \frac{1}{\sqrt{2}}(|\alpha\beta\rangle + |\beta\alpha\rangle)$. Note that the molecular orbitals localized to a spin site can be constructed by mixing the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO) of the RHF canonical orbitals. The BS wave function $|\alpha\beta\rangle$ is described by a linear combination of the singlet ground state and the triplet excited state. The energy gap between two states becomes small when the bond dissociation occurs, but introduction of the spin penalty term $\hat{\mathbf{S}}^2$ to the Hamiltonian can enlarge the gap, as shown in Fig. 3.3b.

3.4 Algorithmic steps

The algorithmic steps described in this section are more tailored to the 2-qubit system. However, the steps remain almost the same for n -qubit system, only differing on initial state preparation, and raising the dimension corresponding to n , among the intermediate terms in the solution.

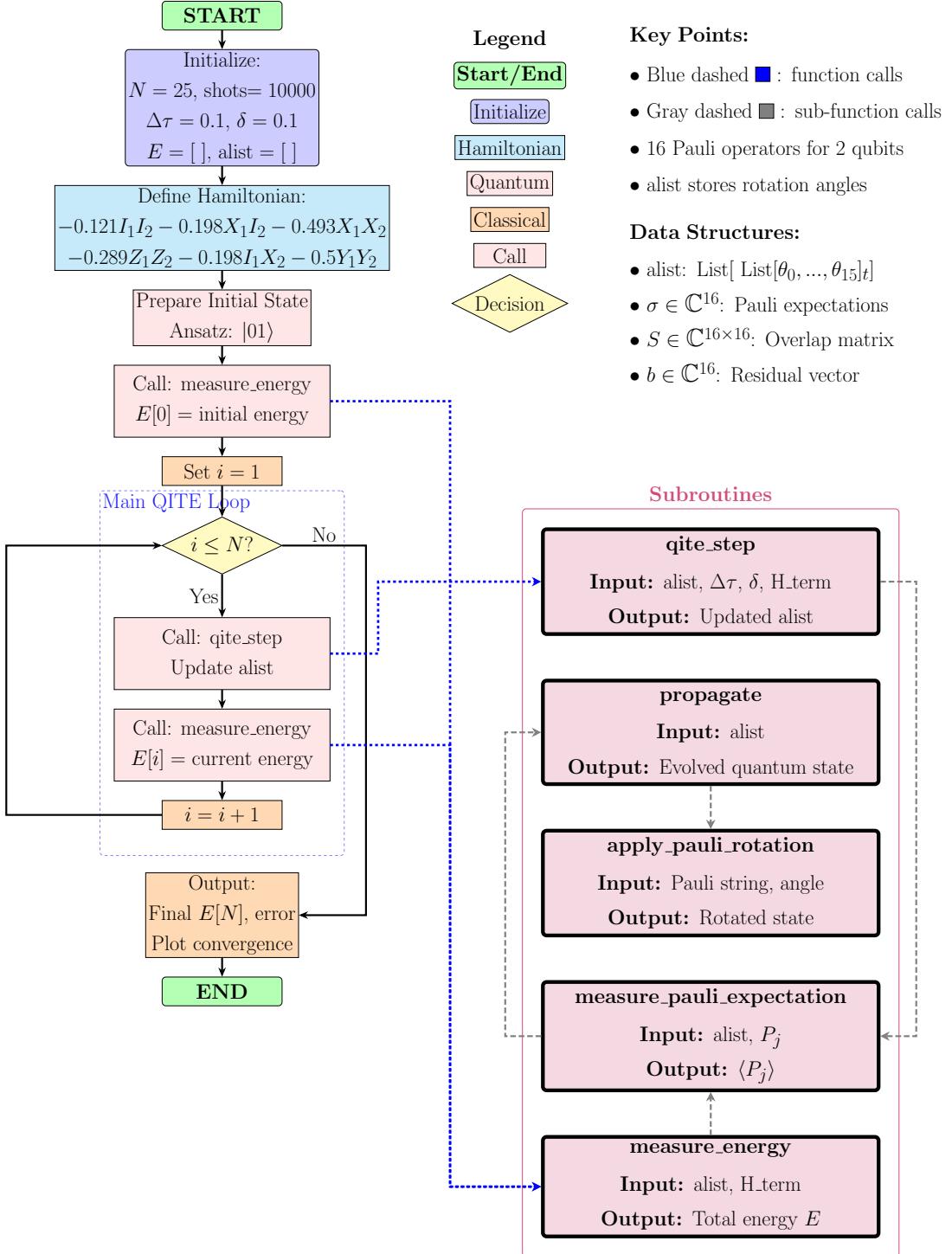


Figure 3.4: Comprehensive flowchart of the 2-qubit Quantum Imaginary Time Evolution (QITE) algorithm. The flowchart shows the complete workflow from initialization through iterative optimization to final energy convergence. Key components include Pauli expectation value measurements, matrix construction for the linear system, and parameter updates through the imaginary time evolution process.

3.4.1 Parameters initialization

Here, we define a Hermitian operator A which contains rotation angles obtained in each iteration of the Trotterized time evolution loop. We initialize A as an empty list at the beginning and continue updating it by solving a linear system.

We run N number of total iterations, each of them evolving with time $\Delta\tau$. We vary $\Delta\tau$ between different systems, tracking the rate of convergence. So, the system evolves for total time, τ equal to N times $\Delta\tau$.

Another very important parameter to take in account while using the shot based sampler for tomography is the regularization parameter, δ . While its introduction is discussed later, we set $\delta = 0.1, 1.0, 10.0, \dots$ as the qubit size increases from 2 qubits by 1.

3.4.2 Hamiltonian construction

The fermionic Hamiltonian operator expressed in equation (3.7), and scaled $\hat{\mathbf{S}}^2$ operator in equation (3.11) are converted to qubit operators using Symmetry-Conserving Bravyi–Kitaev fermion-to-qubit mapping Transformation (SCBKT) [25] and added to obtain our working Hamiltonian,

$$\hat{H}' = SCBKT[\hat{H}] + SCBKT[c\hat{\mathbf{S}}^2]. \quad (3.13)$$

Here, the scaling factor c adjusts the strength of $\hat{\mathbf{S}}^2$. One should make sure that both \hat{H} and $\hat{\mathbf{S}}^2$ act on the same Fock space, with same number of spin orbitals. Now on, for the convenience of notation, we represent our working Hamiltonian \hat{H}' by \hat{H} .

This Hamiltonian is expressed as a linear combination of product of Pauli operators as

$$\hat{H} = \sum_k h_k P_k; \quad h_k \in \mathbb{R}, P^\dagger P = I, \quad (3.14)$$

where h_k is the coefficient corresponding to k^{th} unitary P_k which is a Pauli string. Here, $\sum_k |h_k| = \Lambda$ is the 1-norm of this Linear Combination of Unitaries (LCU) form.

3.4.3 Initial state preparation

The comparison of the electron spin structure for P4 is shown in figure (3.5). If we take the block-spin arrangement, and do JWT transformation where 1 and 0 represent the presence and absence of electron respectively, the RHF and BS states become $|11001100\rangle$ and $|11001010\rangle$ respectively. We however map the fermionic structure using SCBKT mapping, where the RHF and BS states become $|100100\rangle$ and $|100110\rangle$ respectively. One can notice the reduction in number of qubits needed to represent the state with a different mapping. SCBKT provides this advantage over Bravyi–Kitaev and Jordan-Wigner transformations. This is discussed in detail in section 2.4.5. As we just observed, the symmetry broken state is a product state, $|100110\rangle$, it can be prepared in a quantum circuit just using a series of X-gates as shown in figure (3.5). This is an easy construction and shows that using a BS state does not add any overhead during the initial state preparation stage, besides the

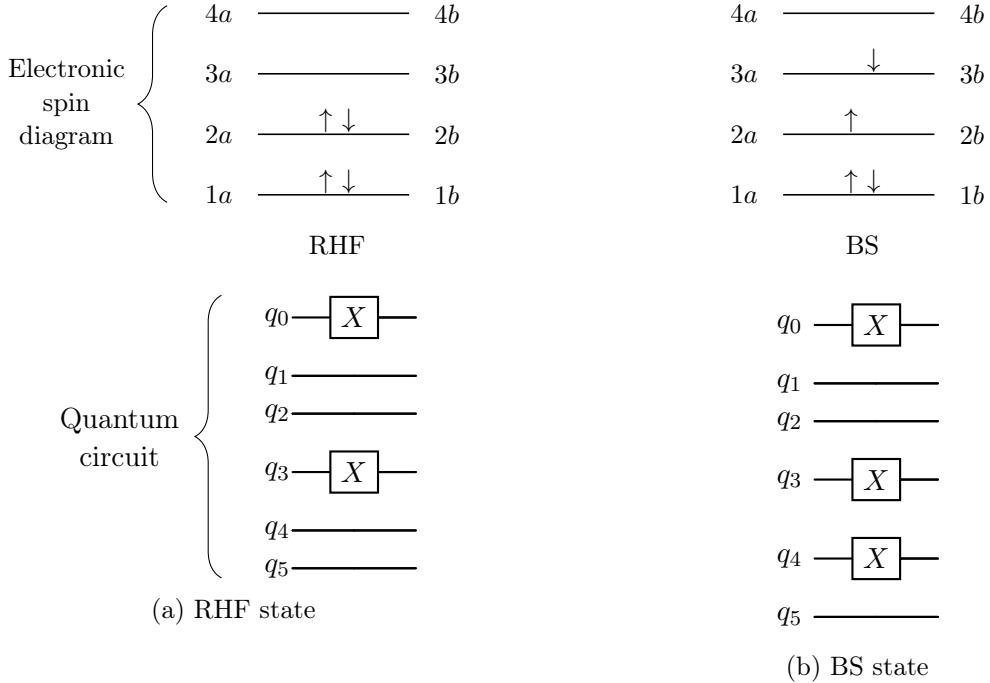


Figure 3.5: Electron spin diagrams and corresponding quantum circuit implementations for RHF and BS states of P4.

information one needs while finding the proper symmetry-broken state off the electronic configuration.

3.4.4 QITE loop

While we have proper definitions of preliminaries as discussed in section 2.6, the solutions are not ready to implement on a quantum computer and have to be simplified for implementation.

As we expect real coefficients in the Hamiltonian, and the state with real elements, b_I in equation (2.136) reduces to $b_I = -2\text{Im}[\langle \Psi | \sigma_I^\dagger | \Delta_0 \rangle]$. Then the normalization constant c is computed as

$$\begin{aligned} c &= \langle \Psi | e^{-2\Delta\tau \hat{H}} | \Psi \rangle \\ &\approx \langle \Psi | (1 - 2\Delta\tau \hat{H}) | \Psi \rangle \quad (\text{1st order expansion}) \\ &= 1 - 2\Delta\tau \langle \Psi | \hat{H} | \Psi \rangle = 1 - 2\Delta\tau \langle \hat{H} \rangle. \end{aligned} \tag{3.15}$$

Similarly, $|\Delta_0\rangle$ can be simplified as

$$\begin{aligned} |\Delta_0\rangle &= \frac{|\bar{\Psi}'\rangle - |\Psi\rangle}{\Delta\tau} \\ &= \frac{c^{\frac{-1}{2}}(|\Psi\rangle - \Delta\tau\hat{H}|\Psi\rangle) - |\Psi\rangle}{\Delta\tau} \\ &= \frac{1}{\Delta\tau} \cdot \left(\frac{1}{\sqrt{c}} - 1 \right) |\Psi\rangle - \frac{\hat{H}}{\sqrt{c}} |\Psi\rangle. \end{aligned} \quad (3.16)$$

S and b from section 2.6 can be computed as the expectation values in the following way:

$$\begin{aligned} b_I &= -2\text{Im}[\langle\Psi|\sigma_I^\dagger|\Delta_0\rangle] \\ &= -2\text{Im} \left[\langle\Psi|\sigma_I^\dagger| \left(\frac{1}{\Delta\tau} \cdot \left(\frac{1}{\sqrt{c}} - 1 \right) |\Psi\rangle - \frac{\hat{H}}{\sqrt{c}} |\Psi\rangle \right) \right] \\ &= -2\text{Im} \left[\frac{1}{\Delta\tau} \cdot \left(\frac{1}{\sqrt{c}} - 1 \right) \langle\Psi|\sigma_I^\dagger|\Psi\rangle - \frac{1}{\sqrt{c}} \langle\Psi|\sigma_I^\dagger\hat{H}|\Psi\rangle \right] \\ &= -2\text{Im} \left[\frac{1}{\Delta\tau} \cdot \left(\frac{1}{\sqrt{c}} - 1 \right) \langle\sigma_I^\dagger\rangle - \frac{1}{\sqrt{c}} \langle\sigma_I^\dagger\hat{H}\rangle \right] \end{aligned} \quad (3.17)$$

$$\text{and } S_{IJ} = \langle\Psi|\sigma_I^\dagger\sigma_J|\Psi\rangle = \langle\sigma_I^\dagger\sigma_J\rangle. \quad (3.18)$$

As summarized in equations (3.15), (3.17) and (3.18), we need the expectation values for various different operators. These are implemented in the intermediate stages of the algorithm within different subroutines.

(1) Subroutines

We describe the entire algorithmic flow based on constituent subroutines. The subroutine in 3.6 iterates over all terms in the Hamiltonian to get the necessary expectation values for the linear system in equation (2.138) and finds necessary rotation angles to propagate to the next state. Total energy in every iteration of the loop is computed by subroutine in 3.7. It iterates over all the terms in the Hamiltonian, (h_k, P_k) and calls another subroutine in 3.8 for each of the Pauli terms. The energy expectation obtained is eventually summed over all terms.

The algorithm requires estimation of the expectation value of each Pauli term within every iteration. Subroutine in figure 3.8 propagates the original ansatz for unitary evolution with the Hamiltonian terms in `alist` and performs measurement on Z-basis. The expectation value is calculated using parity of the Pauli term and counts from many measurement shots.

As described in subroutine in figure 3.9, `propagate` does circuit propagation over all the Pauli rotations. These Pauli rotations are achieved by subroutine in figure 3.10 which constructs the corresponding unitary $e^{-i\theta\frac{P}{2}}$ upon basis conversion with a RZ and Controlled-NOT gates. Representative circuits in figure 3.10 show how each Pauli term and their corresponding coefficients are applied in a circuit in the most naive way.

Subroutine 1: qite_step

Input: `alist`, $\Delta\tau$, `delta`, `H_term` (Hamiltonian terms)

Purpose: Perform one QITE step

For each Hamiltonian term (P_k, h_k) in `H_term`:

1. Measure Pauli expectations: $\sigma[i] = \langle P_i \rangle$ for $i = 0$ to 15
2. Compute overlap matrix: $S_{ij} = \langle P_i P_j \rangle$
3. Compute normalization: $c = \sqrt{|1 - 2\Delta\tau \cdot h_k \langle P_k \rangle|}$
4. Compute residual: $b_i = \Im \left[\frac{\langle P_i \rangle / c - \langle P_i \rangle}{\Delta\tau} - h_k \langle P_i P_k \rangle / c \right]$
5. Solve linear system: $(S + S^T + \text{delta} \cdot I)x = -b$
6. Convert to rotation angles: $\theta_i = -2 \cdot \Delta\tau \cdot x_i$ for $i = 0$ to 15
7. Append to `alist`: `alist.append([\theta_0, \theta_1, ..., \theta_{15}])`

Output: Updated `alist` with new rotation angles (one set per Hamiltonian term)

Figure 3.6: Subroutine for one QITE step that updates rotation angles for all Hamiltonian terms.

Subroutine 2: measure_energy

Input: `alist`, `H_term` (Hamiltonian terms)

1. Initialize total energy $E = 0$
2. For each term (h_k, P_k) in `H_term`:
 - Call `measure_pauli_expectation` for Pauli string P_k
 - Obtain $\langle P_k \rangle$
 - Update accumulated energy: $E \leftarrow E + h_k \langle P_k \rangle$
3. Return the total energy E

Output: Total measured energy E

Figure 3.7: Subroutine for the computation of the total system energy by summing the expectation values of all Hamiltonian Pauli terms.

Subroutine 3: measure_pauli_expectation

Input: Pauli string P_j (e.g., XX, YZ)

1. If $P_j = II$: return 1
2. Build quantum circuit
3. Apply ansatz: X on qubit 1 ($|01\rangle$)
4. Call `propagate(alist)`:
 - For each Pauli string rotation
5. Basis rotations for both qubits:
 - If $P_j[q] = X$: apply H
 - If $P_j[q] = Y$: apply $S^\dagger H$
 - If $P_j[q] = Z$: no rotation
6. Measure both qubits
7. Compute the expectation value as

$$\langle P_j \rangle = \sum (-1)^{\text{parity}} \frac{\text{count}}{\text{total}}$$

Output: $\langle P_j \rangle$

Figure 3.8: Subroutine for the evaluation of expectation value of a given Pauli string.

Subroutine 4: propagate

Input: `alist` (list of 16 rotation angles per timestep)

Purpose: Apply the time-evolution operator

1. If `alist` is empty, return (initial state)
2. For each timestep t in `alist`:
 - For each of the 16 Pauli strings P_i :
 - Get angle $\theta_i = \text{alist}[t][i]$
 - If $|\theta_i| > 10^{-10}$ and $P_i \neq II$:
Call `apply_pauli_rotation(P_i, θ_i)`

Effect: Applies $\prod_i e^{-i\theta_i P_i/2}$

Output: Updated quantum state after time-evolution

Figure 3.9: Subroutine for the propagation of quantum state by sequential application of Pauli rotations.

Subroutine 5: apply_pauli_rotation

Input: Pauli string P , rotation angle θ

1. Basis conversion for each qubit:
 - If $P[q] = X$: apply H to qubit q
 - If $P[q] = Y$: apply RX($\pi/2$) to qubit q
2. Apply rotation according to active qubits:
 - 1 active qubit: RZ(θ)
 - 2 active qubits: CNOT-RZ(θ)-CNOT
3. Reverse basis conversion:
 - If $P[q] = X$: apply H to qubit q
 - If $P[q] = Y$: apply RX($-\pi/2$) to qubit q

Effect: Applies $e^{-i\theta P/2}$

Output: Updated circuit with applied Pauli rotation

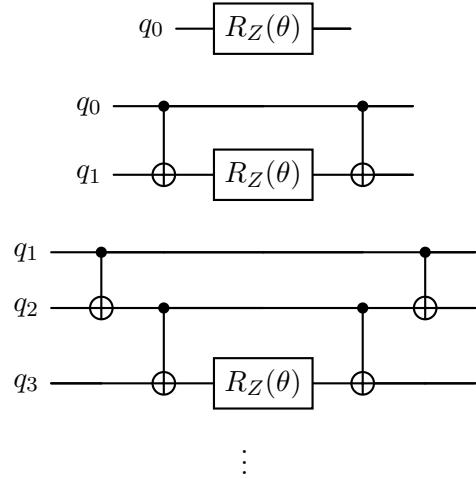


Figure 3.10: Subroutine and its circuit realizations for applying Pauli rotation to qubits.

(2) Solution of linear equation

As mentioned in the subroutine 3.6, we need to solve a linear system, which deserves a critical attention as it is a classical problem and the dimension of matrices grows exponentially on the number of qubits as 4^n ,

$$(S + S^T + \delta \cdot I)x = -b, \quad (3.19)$$

where δ is the regularization parameter. Due to the shot noise, off-diagonal elements in S become noisy with increasing basis size for larger Hamiltonian and the linear system becomes numerically unstable. δ compensates for the shot noise during measurement and prevents the solver from taking the directions dominated by those noise, away from the ground state.

The construction of S matrix needs the expectation values of product of all Pauli terms. Finding b vector needs the expectation values of Paulis, in addition with the product of every term in the Hamiltonian with the Paulis in basis of consideration. Trivially, we take the entire Pauli basis for these expectation values and for constructing b and S . However, as discussed in [8], the component in $b_I \approx 0$ unless σ_I^\dagger contains odd number of Pauli-Y terms. Thus, we can disregard those Pauli strings with even or no Pauli-Y terms, which reduces the number of basis states for consideration. This reduces the size of matrices by almost half the previous size, $2^n \frac{2^n - 1}{2}$. This entire process of finding proper approximation for \hat{A} of equation (2.127) is the process of tomography where we try constructing a state with non-unitary evolution using operator obtained from `qite_step`.

Kasra et al. [46] have proposed an Adiabatic-Quantum Imaginary Time Evolution (A-QITE) method to bypass this state tomography, requiring no measurements on the system. This procedure enables an exponential decrease in the infidelity with the desired Hamiltonian for an adiabatic Hamiltonian for short time of evolution. However, the authors have reported a rapid increase in the cost as time adds up, needing heuristics to control non locality in the adiabatic Hamiltonian.

While solving the linear system has been a classical bottleneck in QITE at this point, using quantum linear solvers [47] could be helpful in the future, sticking with the process of state tomography.

3.5 Computational conditions

We implemented QITE algorithm and demonstrated the advantage of using BS wave function as the initial wave function with penalty introduced in the fermionic Hamiltonian using exact classical emulation. RHF and BS-UHF calculations were carried out using **GAMESS-US** software [48]. The BS-UHF wave function exhibits distinct spatial distributions for spin- α and spin- β orbitals, making its direct application in quantum computation non-trivial. In this work, we generated natural orbitals by diagonalizing the one-electron reduced density matrix obtained from the BS-UHF wave function. Subsequently, molecular orbitals localized to the spin site were manually constructed by mixing the HONO+ i and LUNO- i orbitals, where HONO+ i and LUNO+ i denote the i -th highest occupied natural orbital and the i -th lowest unoccupied natural orbital, respectively. One and two electron atomic orbital integrals computed with **GAMESS-US** were transformed to molecular orbital integrals using our own Python3 code, and the fermionic Hamiltonian was constructed with **OpenFermion** [49] package.

We focused on three molecular systems: H₂ molecule, N₂ molecule, and P4 cluster. The study of bond dissociation of a N₂ molecule in this work is limited to ITE with matrix exponentiation only, to verify that our work also achieves advantage for larger molecules. For N₂ molecule, STO-3G basis set was used with six electrons in twelve spin orbitals. Statevector based numerical simulations were performed for the H₂ molecule and a P4 cluster with QITE. In the study of the ground state energy of the H₂ molecule under a single bond dissociation, we used STO-6G basis set with two electrons in four spin orbitals. In case of a P4 cluster, STO-3G basis set was used with four electrons in eight spin orbitals for the dissociation of intermolecular hydrogen bonds. To perform fermion-to-qubit mappings, Symmetry-Conserving Bravyi–Kitaev Transformation (SCBKT) [25] was used. 1500 total iteration steps were taken for the iterative calculation of $\hat{A}[m]$ for H₂ and P4 cluster with time step size, $\Delta\tau = 0.01$. The domain sizes for H₂ and P4 cluster were taken to be 2 and 6 respectively. The strength of the penalty term in (3.12) was set to $c = 1$.

In addition to statevector based simulation, gate based circuits with sampler based simulations were also carried out for H₂ and P4 systems. The shots taken were adjusted

between 2000 and 10000 for all measurements with $\Delta\tau$ being 0.2 and 5.0 respectively. We set $\delta = 0.1$ and 1500.0 for H₂ and P4 simulations respectively. H₂ was simulated on `ibm_kingston` quantum computer without any external noise mitigation techniques.

Chapter 4

Results and discussion

Parts of this chapter are reproduced with permission from Poudel *et al.* [12].

4.1 ITE with matrix exponentiation

We have investigated the time of convergence to the ground state and the fidelity against CAS-CI energy for N₂ molecule under different bond lengths. Simulations were carried out for different ranges of bond length of N–N atoms, RHF (1 Å - 2.9 Å), BS2 (1.6 Å - 3 Å), BS3 (2.1 Å - 3 Å). BS3 converged to BS2 when the bond length was shorter than 2.1 Å, and BS2 calculations for the N₂ molecule with bond length shorter than 1.6 Å converged to the RHF solution, since N₂ molecule exhibits very small open shell character in the range below them. We define the convergence time as the time required to achieve chemical accuracy, when the energy difference between the ITE state and the CAS-CI state, plus $\langle \hat{S}^2 \rangle$ is $\Delta E + \langle \hat{S}^2 \rangle = E_{\text{ITE}} - E_{\text{CAS-CI}} + \langle \hat{S}^2 \rangle \leq 1.0 \text{ kcal mol}^{-1}$. Note that this gives an upper bound on the convergence time because $\langle \hat{S}^2 \rangle \geq 0$. Here, a single time step corresponding to one iteration is taken as 0.1. So, the time used for describing convergence in N₂ molecule corresponds to the number of iterations, each of size 0.1.

Fig. 4.1a shows faster convergence towards the ground state energy for BS2 and BS3 than RHF inspired ITE for a N–N bond length of 2.7 Å. Here, the energy expectation values in case of BS states are calculated with respect to the penalty inspired Hamiltonian. So, the initial energy differences for BS states are larger than the RHF state, while it is not always obvious as long as the coefficients introduced by the penalty term don't make significant impact on the energy estimation which is the case near the region of ground state. Thus, we can understand that the BS states converge even faster than as shown in the figure in 4.1a. In addition to this, an important observation to note is, that the initial fidelity of both the BS2 (≈ 0.14) and BS3 (≈ 0.25) wave functions with respect to the CAS-CI wave function at 2.7 Å bond length is larger than that of RHF wave function (≈ 0.08). The initial fidelity of the BS2 and BS3 wave functions in the bond dissociation limit is ≈ 0.12 and ≈ 0.25 respectively. The contribution of the penalty term in (3.12) is shown in

the lower panel of Fig. 4.1a. For the RHF wave function, $\langle \hat{\mathbf{S}}^2 \rangle$ remains zero throughout the evolution, and it shows slow convergence towards the ground state. Interestingly, when the initial state is a BS wave function, the penalty contributes significantly only in the early stage of the evolution. This brief but large contribution ensures the initial state quickly evolves towards the ground state.

Next, we investigate whether the BS wave function is a suitable initial state for all bond lengths. Fig. 4.1b shows that this is not always the case. For shorter bond lengths (blue region), the RHF converges faster than both BS2 and BS3. Note that for shorter bond lengths, the BS-UHF calculation converged to the RHF solution instead of the open shell state. At these geometries, the BS2 and BS3 wave functions were generated by manually mixing the occupied and unoccupied orbitals of the RHF wave function, followed by placing the unpaired electrons into the localized orbitals. Consequently, the BS2 and BS3 wave functions, used as the initial wave functions for BS-QITE, have higher energies than the RHF wave function. As the bond length increases, the convergence time for BS wave functions decreases, while the performance of the RHF wave function worsens. For intermediate bond lengths (yellow region), the BS wave functions outperform the RHF wave function, with BS2 performing slightly better than BS3. Finally, we observe that for long bond lengths (green region), the BS3 wave function is far more suitable initial state than RHF or even BS2. It is also important to note here that we used the knowledge of convergence time versus interatomic distance for the range described above for the case of BS2 and BS3 to extrapolate the plots for bond lengths ranging from 1.0 Å to 3.0 Å in Fig. 4.1b.

In order to determine when a particular wave function is most suitable as the initial state, we compute the diradical character as a function of bond length. Diradical character is a measure of the open shell character of the wave function, and it can be calculated at the spin-projected spin-unrestricted HF (PUHF) level [50],

$$y_i^{\text{PUHF}} = 1 - \frac{2(1 - n_{\text{LUNO}+i})}{1 + (1 - n_{\text{LUNO}+i})^2}. \quad (4.1)$$

Here, $n_{\text{LUNO}+i}$ is the occupation number of the i -th lowest unoccupied natural orbital (LUNO+ i). For convenience, we denote the diradical character y_i^{PUHF} as y throughout the rest of this thesis. It ranges from 0 to 1, a molecule with $y > 0$ is considered to have open shell character which becomes more prominent as y increases. At the extremes, $y = 0$ for a closed shell molecule, and $y = 1$ for a pure diradical.

For N₂ bond dissociation, we focus on two diradical characters y_π and y_σ for open shell characters of valence π and σ bonds, respectively, as shown in Fig. 4.1c. Numerical simulations reveal that ITE with BS2 converges as fast as the conventional RHF-reference ITE within the bond length of 1.62 Å corresponding to the diradical character $y_\pi = 0.48$. For further dissociation, BS2 wave function performs better until the bond length reaches 2.08 Å with $y_\sigma = 0.35$. Beyond this bond length, BS3 is the most suitable to use as a starting wave function.

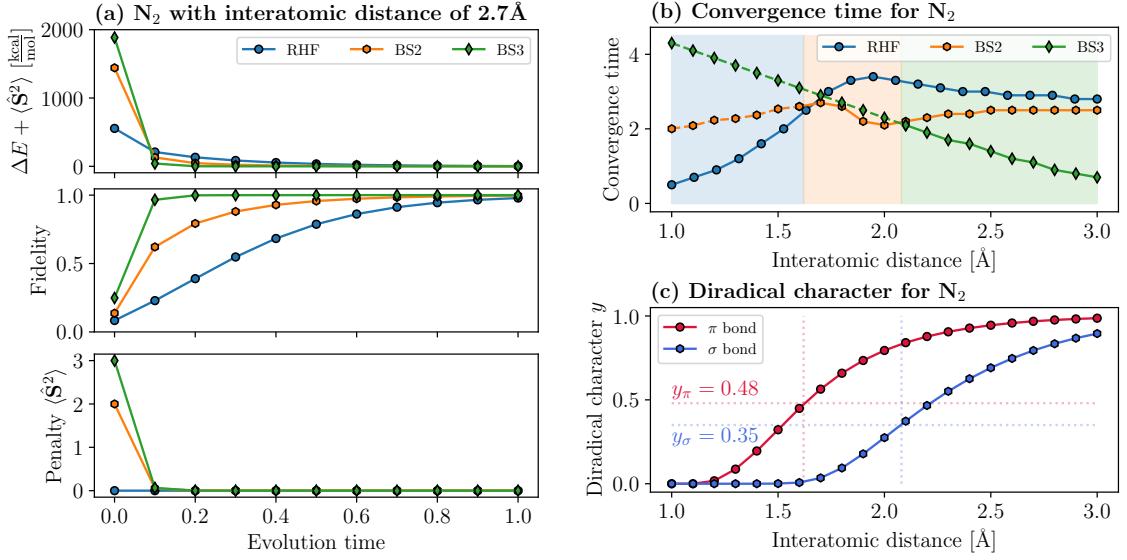


Figure 4.1: (a) Comparison of the convergence behavior of ITE in N_2 molecule with the RHF, BS2, and BS3 wave functions as the starting wave function. Top panel shows the energy difference $\Delta E + \langle \hat{\mathbf{S}}^2 \rangle$ during the evolution. Fidelity of ITE evolved wave function with respect to the CAS-CI wave function in the middle panel, and contribution of the penalty term $\langle \hat{\mathbf{S}}^2 \rangle$ in the bottom panel. (b) Time of convergence to attain the chemical accuracy $\Delta E + \langle \hat{\mathbf{S}}^2 \rangle \leq 1.0 \text{ kcal mol}^{-1}$ for bond length from 1.0 Å to 3.0 Å. Background colors show the recommended region of wave function as starting wave function depending on the N–N bond length. (c) Diradical character y at the region of bond dissociation. The points of intersection of red and blue lines with corresponding red and blue curves indicate the recommended points to change the initial wave functions described in (b).

However, it remains an open problem to use the diradical character to determine the proper starting wave function without the aid of these simulation works.

4.2 Simulator results

4.2.1 Numerical simulation with statevector

We now focus on the H_2 molecule and get into the details of the behavior of QITE throughout its PES. As discussed earlier, the maximum spin multiplicity involved in the BS wave function for a H_2 molecule is $S = 1$. This leads to an extra penalty contribution of $S(S + 1) = 2$ to the energy of the excited triplet, leading to an increase in the energy gap with respect to the ground state.

Fig. 4.2a shows the difference in ground state energy with two different wave function driven QITE with CAS-CI values and the comparison of fidelities. The chemical accuracy

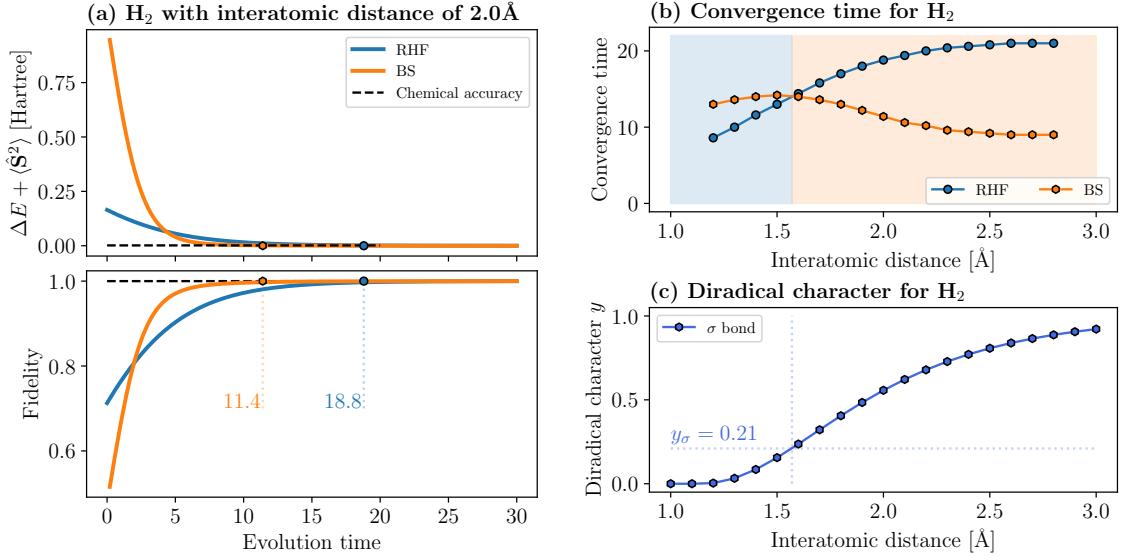


Figure 4.2: (a) Comparison of the convergence behavior of QITE in H_2 molecule with the RHF and BS wave functions as the starting wave function. Top panel shows the energy difference between the ground state energy calculated using CAS-CI and QITE with unitary approximation plus $\langle \hat{S}^2 \rangle$. Dashed line represents chemical accuracy given by 1.59×10^{-3} Hartree. Bottom panel shows fidelity of QITE evolved wave function with respect to the CAS-CI wave function. Dashed line marks unit fidelity. (b) Recommended region of wave function as starting wave function depending on the bond length between hydrogen atoms. (c) Diradical character y at the region of bond dissociation. The point of intersection of two straight lines with the curve indicate the recommended point to change the initial wave function described in (b).

could be attained within 260 steps using BS wave function, while it took 440 steps with HF wave function for a bond length of 2.0\AA with diradical character $y = 0.56$. This simulation was carried out with domain size of 2 considering two electrons in the Hamiltonian. Here we discard the concept of locality while approximating the unitary operator from a nonunitary time evolution.

Similar to what we observed during matrix exponentiation of the N_2 molecule, it is not wise to use a BS wave function in all the context of H_2 molecule too. Around equilibrium, at bond length of around 0.7\AA and up to 1.56\AA , the HF wave function is a better choice for starting wave function. Bond length of 1.56\AA corresponds to the diradical character $y = 0.21$. For H_2 beyond this diradical character, BS wave function is the better choice. Fig. 4.2b shows the time required to converge to the ground state for QITE under various bond lengths. Here, ‘‘converge’’ means that the QITE achieves chemical accuracy ($\Delta E + \langle \hat{S}^2 \rangle = E_{\text{QITE}} - E_{\text{CAS-CI}} + \langle \hat{S}^2 \rangle \leq 1.0 \text{ kcal mol}^{-1}$ (1.59×10^{-3} Hartree))). QITE with RHF wave function takes a long time to converge as the bond length increases, while it is

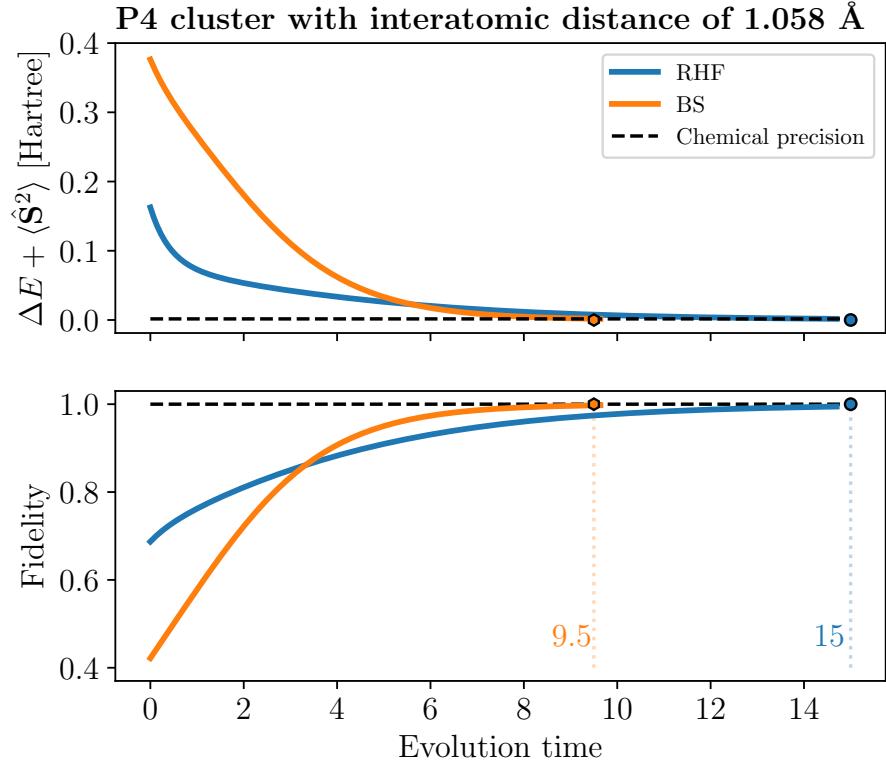


Figure 4.3: Comparison of the energy difference between the ground state energy calculated using CAS-CI and using HF and BS wave function inspired QITE with unitary approximation plus $\langle \hat{S}^2 \rangle$ in the top panel. Bottom panel shows the fidelity of QITE evolved wave function with respect to the CAS-CI wave function for tetrahydrogen cluster with H–H interatomic distance of $\approx 1.058 \text{ \AA}$.

the opposite in the case of QITE with BS wave function. The background colors indicate recommended choice of starting wave function under varying bond lengths.

The diradical character y of the H_2 molecule during bond dissociation is shown in Fig. 4.2c. BS-UHF converged to the RHF solution and y is calculated to be 0 for bond length up to 1.2 \AA . Increasing the bond length further results in an increase of the diradical character, motivating the molecular geometry to have unpaired electrons in different orbitals. At a bond length of around 3.0 \AA , the diradical character is very close to 1. This is the region where the energy gap between the ground state and the first excited state is close to 0.

Tetrahydrogen cluster shows a very strong open shell character at the square geometry with the bond length of $\approx 1.058 \text{ \AA}$. In this geometry, even considering the domain size of 6 qubits, the HF wave function assisted QITE takes longer time to converge to the ground state. Fig. 4.3 confirms the faster convergence of BS wave function assisted QITE

with the same domain size during numerical simulation. It compares the fidelities of the wave function with BS and HF assisted QITE with CAS-CI wave function during the time evolution. Though HF assisted QITE has higher fidelity as the evolution initiates, BS assisted QITE rises above it and achieves chemical precision early. The chemical precision was attained in ≈ 950 iteration steps with BS assisted QITE, while it took ≈ 1500 steps with HF assisted QITE. The conditions taken for the simulation are described in Section 3.5.

4.2.2 Simulation with sampler

While we ran the simulations using statevector for mid circuit expectations before, we also ran realistic yet noise-less simulations using sampler for intermediate expectation values calculation for many shots using Qiskit’s `AerSimulator` [14], version 2.2.3. The algorithm for implementing sampler is discussed in section 3.4. Sampler based simulations were done for both H₂ and P4 systems. Figure 4.4 shows the results for sampler based simulations

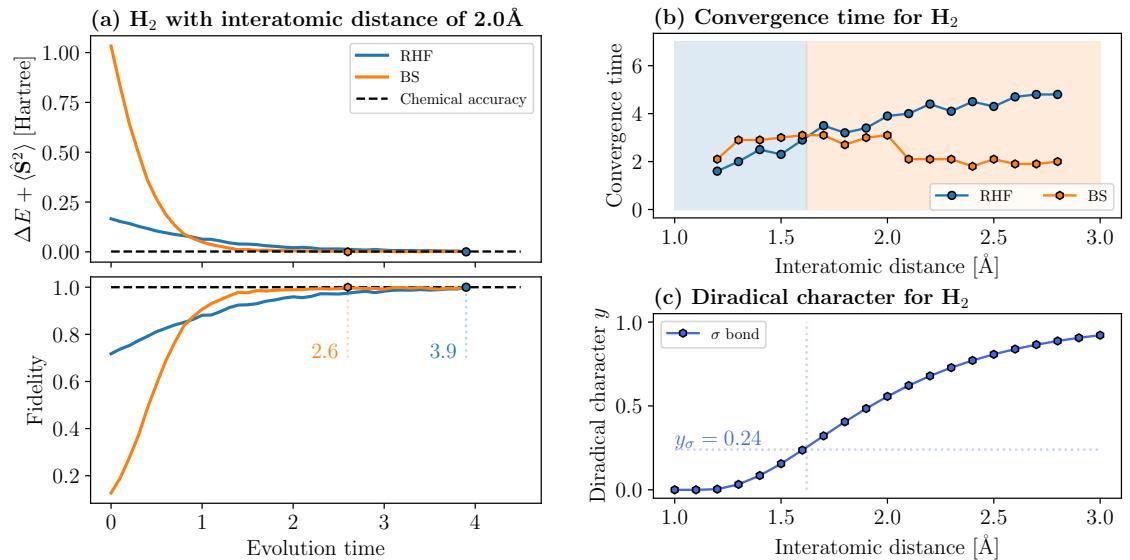


Figure 4.4: (a) Comparison of the convergence behavior of QITE in H₂ molecule with the RHF and BS wave functions as the starting wave function using sampler. Top panel shows the energy difference between the ground state energy calculated using CAS-CI and QITE plus $\langle \hat{S}^2 \rangle$. Dashed line represents chemical accuracy given by 1.59×10^{-3} Hartree. Bottom panel shows fidelity of QITE evolved wave function with respect to the CAS-CI wave function. Dashed line marks unit fidelity. (b) Recommended region of wave function as starting wave function depending on the bond length between hydrogen atoms. (c) Diradical character y at the region of bond dissociation. The point of intersection of two straight lines with the curve indicate the recommended point to change the initial wave function described in (b).

of QITE with step size of 0.1 and 1024 shots of every measurements. The evolution time is measured in atomic units as before. Since the shot based measurements are stochastic, the times of convergence towards the ground state by chemical accuracy of 1.59×10^{-3} Hartree is near to the statevector based simulation, though not exactly the same. For the interatomic distance of 2.0 Å, BS wave function motivated QITE with sampler converged in exactly the same number of QITE steps as observed with statevector. However, RHF based QITE was seen converging a bit earlier than previous statevector instance. This difference is due to shot errors which can be eliminated by increasing the number of shots. For the ease of simulations in our case, we use regularizer in a way discussed in section 3.4 to account for stabilizing the time evolution trajectory distorted due to shot errors. For simulating H₂, we set the regularizer to be 0.1. Higher values of this term allow for better solutions, yet we settle on averagely performing values to make the simulation more natural.

Results show that RHF performs better until the H₂ geometry is $\approx 1.62\text{\AA}$, with diradical character 0.24, and BS is the best choice beyond that. This is close to the numbers obtained from statevector based simulations (pivoting geometry $\approx 1.56\text{\AA}$, diradical character = 0.21).

Simulations were also carried out for P4 cluster using sampler based measurements for state tomography. Sampler based simulations for larger system like P4 require large number of measurement circuits with higher measurement shots. In addition, the dimension of Hamiltonian, **alist**, **b** and **S** grow exponentially as 4^n for n-qubits. This requires solving the system of linear form in equation 3.19 with dimension of 4^n and at the same time measuring the expectations of the permutation of all n-term Paulis with respect to the quantum state.

As talked about in section (2), we do not need to obtain the expectation values for all Pauli terms, but just those terms with odd number of Y-operators in them. While this reduces the dimension of solution to $2^n \frac{2^n - 1}{2}$, measuring the expectations of terms in the Hamiltonian with respect to $2^n \frac{2^n - 1}{2}$ terms is still computationally slow and we need that many copies of the state. We optimize this computation and simulate P4 cluster within a computationally doable region for a personal computer.

4.3 Optimizing tomography for simulating larger systems

We optimize the overall simulation, both for RHF and BS inspired QITE on the tomographic step. For the general case, expectations for all constituting unitaries in the form of Paulis in Hamiltonian are obtained separately with respect to the quantum state. This causes enormous increase in the number of measurements.

First of all, we eliminate very small rotations, $< 10^{-10}$ while applying gate operations. To optimize the number of measurements needed, we group the commuting terms in the Hamiltonian together and obtain expectation values of those commuting terms together with respect to the Paulis with odd number of Pauli- Y operators. One needs to understand that, for a typical example of P4, the dimensions of terms **b**, **S**, **alist** are ~ 2100 during

Algorithm: Optimized QITE with Hamiltonian Grouping

Input: RHF Hamiltonian H with 85 terms (bond length $\approx 1.058 \text{ \AA}$), time step $\Delta\tau$, steps N , regularization δ , shots

Purpose: Find ground state energy through quantum imaginary time evolution

1. Initialize: $E = \text{array of size } (N + 1)$, $\text{alist} = []$
2. **Group Hamiltonian terms by commutativity:**
 - Apply greedy grouping algorithm to 85 terms
 - Obtain G commuting groups (typically $G \approx 6$ for RHF-P4)
3. Measure initial energy: $E[0] = \text{measure_energy}(\text{alist}, H)$
4. For $i = 1$ to N :
 - **For** $j = 1$ to G :
 - Call $\text{process_group}(\text{alist}, \text{group}[j], \Delta\tau, \delta)$
 - Update alist with new rotation angles
 - Measure energy: $E[i] = \text{measure_energy}(\text{alist}, H)$
 - If $|E[i] - E[i - 1]| < \epsilon$: convergence achieved, break
5. Return energy trajectory $E[0 : i]$

Output: Energy array E showing convergence to ground state

Optimization: Reduces expectation measurements from 85 to G per QITE step ($\sim 7\times$ speedup)

Figure 4.5: Optimized QITE algorithm with grouped Hamiltonian terms.

the solution and `alist` containing angles is later expanded to 4096 terms with sparsity while finding the energy expectation value. The optimized implementation is described in figure 4.5 and operation within a single commuting group is discussed as a subroutine in figure 4.6.

In figure 4.5, the steps 2 and 4 in the algorithm with bold font are integral for optimizing the tomography process.

Subroutine: process_group

Input: `alist`, `group[j]` (commuting Hamiltonian terms), $\Delta\tau$, δ

Purpose: Process all commuting terms in group j together

1. Measure Pauli expectations (once per group):
 - Call `measure_pauli_expectation` for each kept basis term
 - Obtain $\sigma[i] = \langle P_i \rangle$ for $i \in \mathcal{I}_{\text{kept}}$ ($K \approx 2100$ terms)
2. Initialize accumulated residual: $b_{\text{total}} = 0 \in \mathbb{C}^K$
3. For each Hamiltonian term (P_k, h_k) in `group[j]`:
 - Compute normalization: $c = \sqrt{|1 - 2\Delta\tau \cdot h_k \langle P_k \rangle|}$
 - Compute residual for this term:

$$b_i = \Im \left[\frac{\langle P_i \rangle / c - \langle P_i \rangle}{\Delta\tau} - h_k \langle P_i P_k \rangle / c \right]$$
 - Accumulate: $b_{\text{total}} \leftarrow b_{\text{total}} + b$
4. Compute overlap matrix (once per group): $S_{ij} = \langle P_i P_j \rangle$
5. Solve linear system: $(S + S^T + \delta \cdot I)x = -b_{\text{total}}$
6. Convert to rotation angles: $\theta_i = -2\Delta\tau \cdot x_i$ for all kept basis indices
7. Append to `alist`: `alist.append([\theta_0, \theta_1, ..., \theta_{4095}])`

Output: Updated `alist` with one new set of rotation angles (for entire group)

Key advantage: σ measured once and S computed once per group, not per term

Figure 4.6: Subroutine for processing one group of commuting Hamiltonian terms efficiently.

Figure 4.7 shows the simulation result of optimized sampler based simulation for P4 cluster with initial HF state. The symmetry conserved state of the P4 cluster at bond geometry of 1.058 Å has 83 constituent 6-qubit Pauli terms in the Hamiltonian. While the fermion-to-qubit mapping is done with SCBKT, the original requirement of 8-qubits based on STO-6G reduces to 6 qubits. Since this is a sampler based simulation, the number of measurement shots and shot noise are very integral in the simulation results. Due to the necessity to reduce the computational overhead, we use only 2000 shots for every measurements, and to compensate for the discrepancies introduced by this little number of shots and noise from shots, we use very high regularization parameter as 1500. The figure shows the simulation curve until the energy difference is ≈ 0.065 . We could not carry out further simulations beyond this, due to the requirement of larger time of simulation. In fact, extrapolation with log transform shows it would need years to reach chemical accuracy.

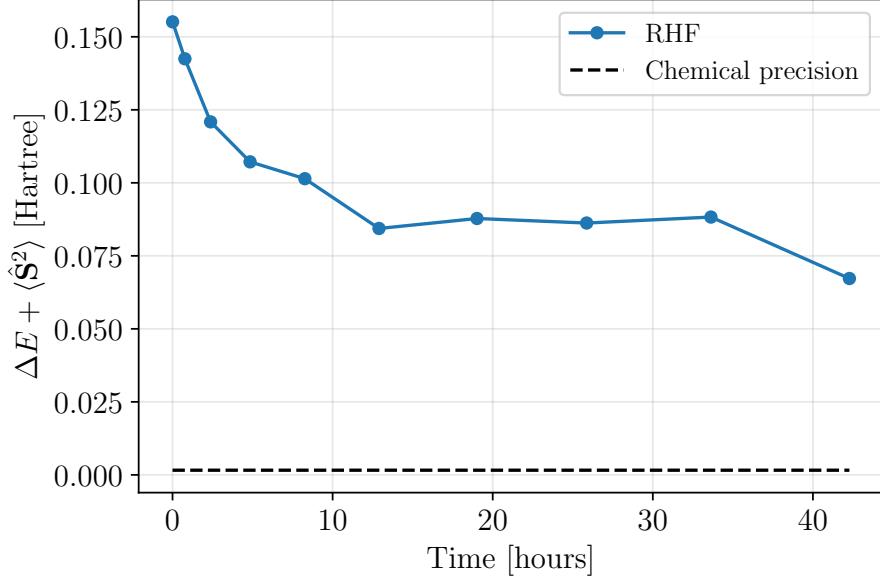


Figure 4.7: QITE implementation of P4 cluster, bond length 1.058 Å with RHF wavefunction by optimizing tomography with Pauli grouping the Hamiltonian terms. Parameters include $\delta = 1500.0$, $\Delta\tau = 5.0$ with 2000 shots for every measurement.

4.4 Results on IBM quantum device

We ran our simulation and verified faster convergence of QITE with symmetry-broken wave function on a IBM quantum machine for H₂ geometry with bond length of 2.7 Å. The simulations were carried out on 18th October, 2025 on `ibm_kingston`. Existing noisy quantum computers are NISQ-era quantum computers. They face the problems of decoherence and noise. The IBM quantum computer we used is a superconducting quantum computer and had the relaxation time and dephasing/coherence time as 337.98 μ s and 168.18 μ s respectively. The relaxation time measures how long a qubit can stay in the excited state before losing energy to move to the ground state. The dephasing/coherence time measures how long a qubit can maintain the phase of its superposition.

For two qubit simulation on a IBM quantum computer for our purpose, we don't need techniques for quantum error correction (QEC) too, which is indispensable for simulating in fault-tolerant era. The simulation was carried out without noise mitigation or suppression techniques in use explicitly. At the time of simulation, median and the best values of two qubits error were 3.83×10^{-3} and 1.17×10^{-3} , respectively. The median of readout error was 1.17×10^{-2} . The arrangement of qubits in the `ibm_kingston` machine is shown in figure 4.9.

The chosen layout selects the physical qubits with the least noise for the best performance. `sabre` layout [51] was used with `optimization_level` set to 3 using qiskit transpiler [14] and time step of 0.2.

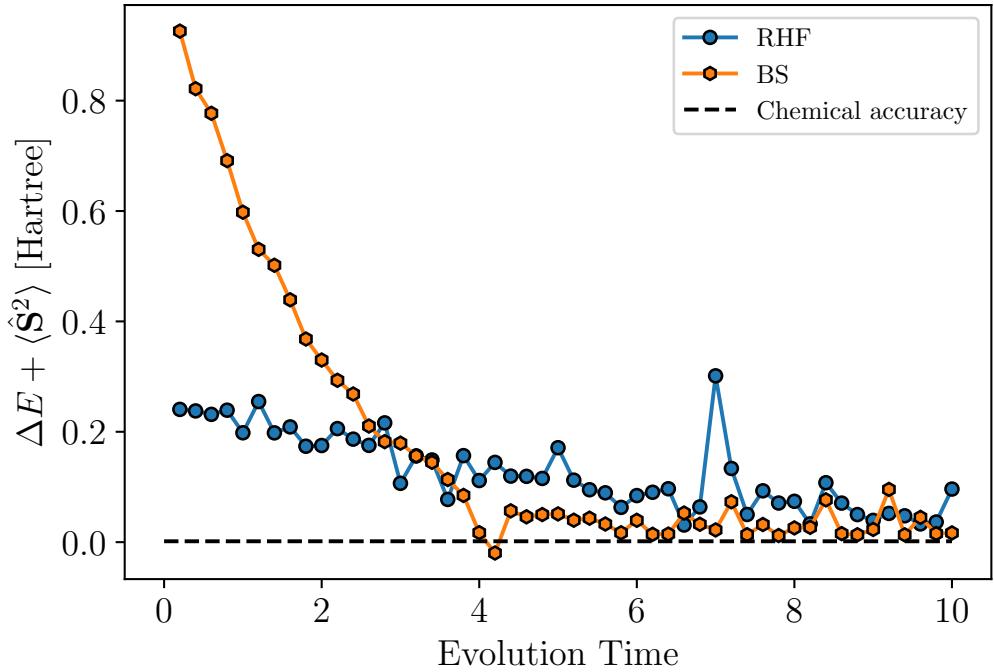


Figure 4.8: QITE comparison of H_2 molecule, bond length 2.7 Å with RHF and BS wavefunctions on `ibm_kingston` with 10000 measurement shots for expectation values.

For H_2 , the circuit depth was ~ 11 for both RHF and BS inspired QITE. Yet, the circuit depth for BS inspired QITE is larger than RHF inspired one for larger correlated systems. However, QITE starting with BS requires less number of iterations to evolve than the next, causing faster convergence than starting with RHF wave function. Figure 4.8 shows that QITE-RHF is still not converging until the evolution time equal to 10 atomic units, while QITE-BS almost attained the chemical accuracy at 4 atomic units. This is a very huge improvement in the time of convergence, and it implies that for the case of smaller correlated systems, starting with BS wave function during QITE is the best choice.

However, one has to be careful to study the overhead of circuit depth due to breaking of symmetry and Pauli-terms with higher coefficients added to the symmetry broken Hamiltonian. The noisy behavior of the curve seen in figure 4.8 is mostly due to shot errors and possibly readout errors and single and 2-qubit errors to some extent. Due to the reduced circuit depth by transpilation and the use of only two qubits, best chosen during mapping based on qubit readout errors and more, within transpilation package, we haven't taken care of decoherence and other noisy factors that exist in the NISQ settings.

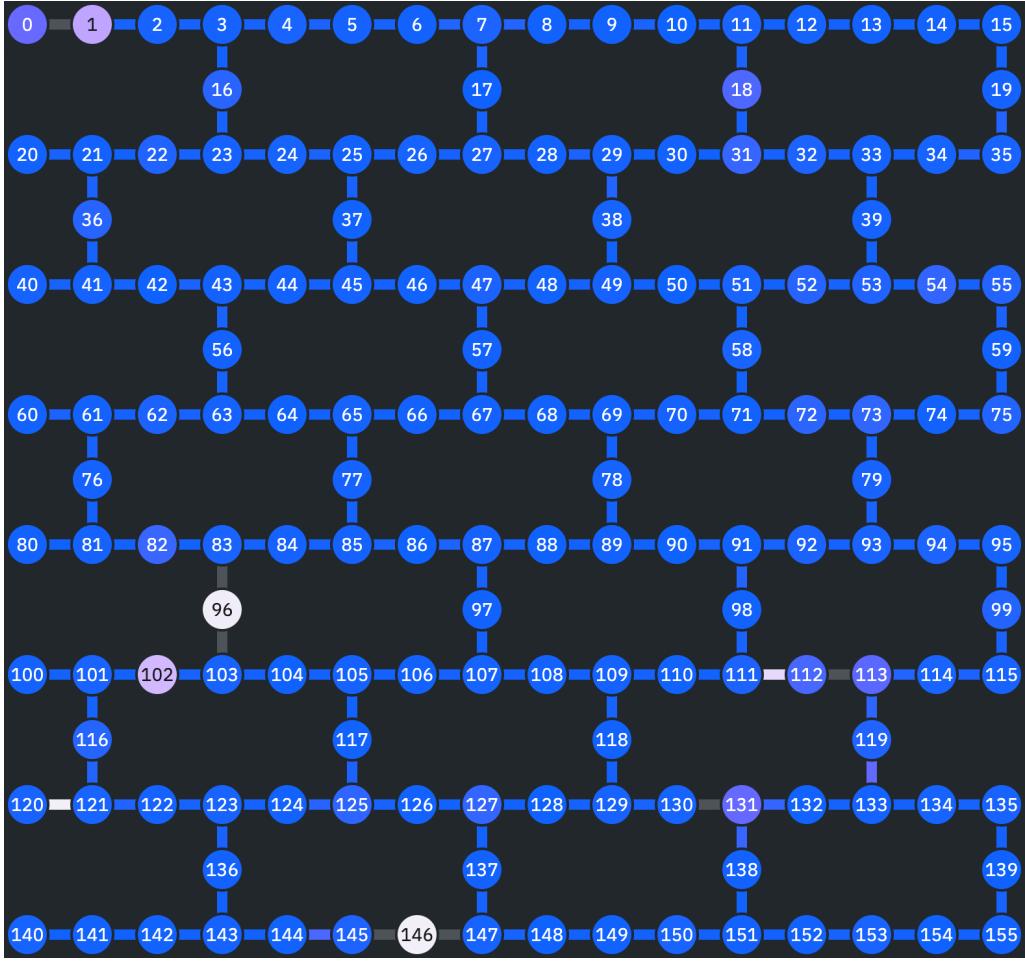


Figure 4.9: Qubit arrangement in `ibm_kingston` Heron r2 processor [52]. Heron r2 processor has 156 total qubits labeled from 0-155. Since these are superconducting qubits, there is no all-to-all connectivity among the qubits and thus need swap gates to perform quantum operations between non-neighboring qubits.

4.5 Initial fidelity in strongly correlated region

The initial fidelity of BS wave functions against the CAS-CI wave function at the corresponding bond length is greater than RHF wave function during the bond dissociation process as shown in Fig. 4.10. We considered the case of a N₂ molecule with six spins to study this behavior. As the molecule approaches bond dissociation limit, the initial fidelity of BS3 wave function approaches 0.25, while that of RHF decreases to 0.05. Thus, the BS wave function is itself a better approximation for open shell system than a RHF wave function. However, for the case of H₂ molecule with two spins and P4 cluster with four spins,

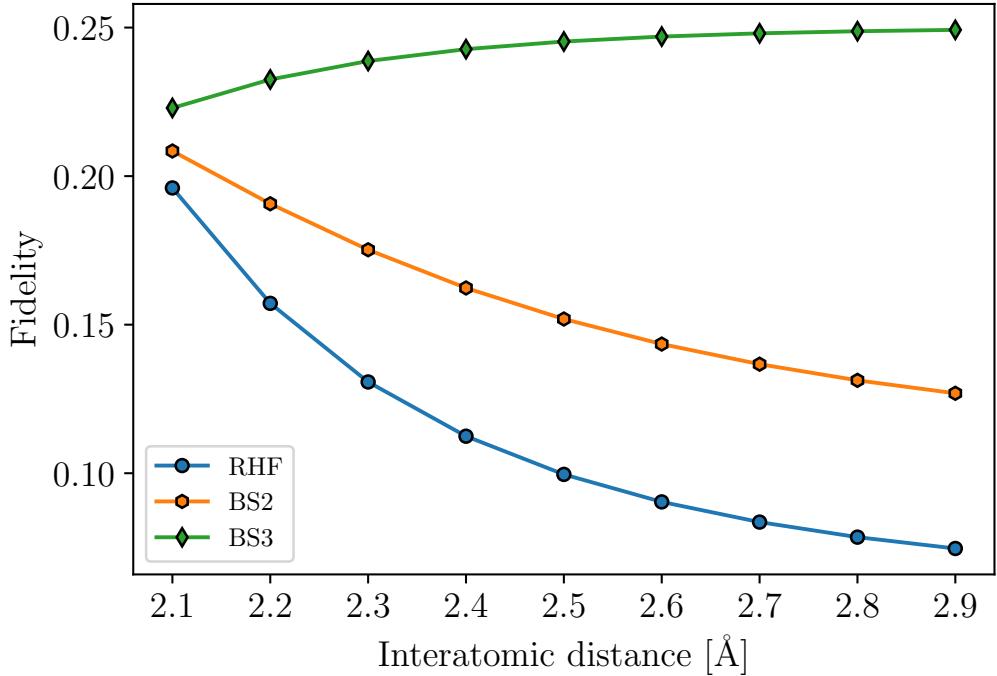


Figure 4.10: Initial fidelities of the wave functions of N_2 molecule against the CAS-CI wave function.

HF wave function was found to have larger initial overlap than BS wave function even near the bond dissociation limit. We can understand that for a larger system executing open shell character, the BS wave function as initial wave function can become superior in cases where the initial overlap of the wave function with the CAS-CI values is crucial.

4.6 QITE algorithmic complexity

Consider n as the number of qubits, the number of terms in the Hamiltonian as T , total number of iterations needed for QITE convergence as N , size of Pauli basis, 4^n as P and number of shots per measurement as S . Let $0 \leq \alpha \leq 1$ be the sparsity factor of `alist`, which quantifies the fraction of angles in `alist` with non-negligible angles. This depends on the structure of the Hamiltonian and the convergence behavior.

Let us make the breakdown of the complexity step by step in the algorithmic pipeline. At step k ,

- The number of time steps stored in `alist` is $k \times T$.
- The number of Pauli strings per time step is 4^n .
- Hence, the total number of active rotations is $kT\alpha 4^n$; where $\alpha \in [0, 1]$ denotes the fraction of active Pauli terms.

- Each Pauli rotation is implemented using CNOT ladders. For m active qubits, the total gate count per rotation is approximately $4m - 1$, consisting of
 - $2(m - 1)$ CNOT gates for forward and backward ladders,
 - 1 single-qubit rotation gate,
 - $\sim 2m$ single-qubit basis rotation gates.
- Therefore, the total number of gates per circuit is approximately $(4m - 1)kT\alpha 4^n$.
- Since expectation values are obtained T times, measuring all 4^n Pauli operators, the total number of measurement circuits is $T \times 4^n$.
- In addition, the function `measure_energy` is called T times, contributing an additional T circuits.
- Consequently, the total number of gates at step k is approximately

$$T(4^n + 1)(4m - 1)kT\alpha 4^n \approx T^2 4^{2n} (4m - 1)k\alpha.$$

Considering N steps are required for converging to the ground state,

$$\begin{aligned} \text{Total gates} &= \sum_{k=1}^N [\text{circuits at step } k] \times [\text{gates per circuit at step } k] \\ &\approx \sum_{k=1}^N T^2 4^{2n} (4m - 1)k\alpha \\ &\approx T^2 4^{2n} (4m - 1)\alpha \sum_{k=1}^N k \\ &\approx T^2 4^{2n} (4m - 1)\alpha \frac{N(N + 1)}{2} \\ &\approx \frac{N^2 T^2 4^{2n} (4m - 1)\alpha}{2}. \end{aligned} \tag{4.2}$$

Considering, Pauli operators with m active qubits, in average we take that the number of X, Y, Z operators is $m/3$ each. Thus, in case where Hadamard gates are needed for X , the number of H gates becomes $\frac{2m}{3}$, and the number of Rx gates becomes $\frac{2m}{3}$ for Y operations. Thus, a single rotation for implementing X and Y need $\frac{2m}{3}$ gates each. Table 4.1 summarizes the complexity of each steps. In addition, gate level breakdown for H, Rx, Rz and CNOT gates is summarized in table 4.2

The $O(T4^n)$ scaling from quantum state tomography, combined with $O(T4^n)$ from the accumulation of `alist` yields the overall $O(T^2 4^{2n})$ quantum gate complexity as included in table 4.1. The rest of the classical bottleneck arises from processing the tomography data to build S matrix ($O(4^{2n})$) and requiring $(NT \times 4^{3n})$ operations to solve $4^n \times 4^n$ linear system.

Due to this exponential squared complexity, we have not done simulations for molecules that require more than 6 qubits in QITE. As shown and discussed in the section 4.3, methods like Pauli grouping can be used to speed up the tomography process, but it does

Component	Depends On	Complexity	Scaling	Description
Time steps in <code>alist</code>	k, T	kT	$O(kT)$	Number of angle sets accumulated until step k
Circuit depth	k, T, n, α, m	$(4m - 1) \cdot kT\alpha \cdot 4^n$	$O(kTm4^n)$	Gates per measurement circuit at step k
Circuits executed	T, n	$T(4^n + 1)$	$O(T4^n)$	Quantum circuits run during step k
Gates	k, T, n, α, m	$T(4^n + 1) \cdot (4m - 1) \cdot kT\alpha \cdot 4^n$	$O(kT^2m4^{2n})$	Total gates in all circuits at step k
Measurement shots	T, n, S	$T(4^n + 1) \cdot S$	$O(T4^nS)$	Total measurements at step k

Table 4.1: Quantum computational complexity at QITE step k . Parameters: k = current step, T = Hamiltonian terms, n = qubits, α = sparsity, m = average active qubits per Pauli, S = shots per circuit.

Gate Type	Count per Rotation	Total at Step k	Scaling	Role
Hadamard (H)	$\frac{2m}{3}$	$\frac{2m}{3}kT^2\alpha4^n(4^n+1)$	$O(mkT^24^{2n})$	X-basis change
Rx rotation	$\frac{2m}{3}$	$\frac{2m}{3}kT^2\alpha4^n(4^n+1)$	$O(mkT^24^{2n})$	Y-basis change
Rz rotation	1	$kT^2\alpha4^n(4^n + 1)$	$O(kT^24^{2n})$	Parametric rotation
CNOT (CX)	$2(m - 1)$	$2(m - 1)kT^2\alpha4^n(4^n + 1)$	$O(mkT^24^{2n})$	Entangling gates
Total	$4m - 1$	$\frac{(4m - 1)kT^2\alpha4^n(4^n + 1)}{3}$	$O(mkT^24^{2n})$	All gates

Table 4.2: Gate type decomposition at QITE step k . The T^2 scaling arises from T Hamiltonian terms requiring $T \times 4^n$ circuits, each with depth proportional to kT .

not bring that high scaling guarantee to the flow. In addition to that, we have mentioned some possible methods to improve the computational complexity in the future works section at 5.2.

In this complexity analysis, we do not account for the diradical character or the energy gap between the ground state and the first excited state, and thus cannot compare the performance of symmetry broken method with the RHF method discussed in this work based on computational complexity. It is necessary to consider that the number of Pauli terms in the symmetry broken Hamiltonian with penalty increases by more than doubles for a bit larger molecules like N₂ compared to the RHF (535/247), which is expected to have even larger difference for molecules beyond this size. In addition to that, the coefficients of the terms also leads those in RHF by a lot. This is a significant limitation to breaking the symmetry and introducing spin penalty. Thus, this might not be the optimal method for simulating larger molecules with QITE. Yet, we cannot confirm the complexity without analyzing the energy gap, diradical character and the convergence speed BS state provides devoid of the size of the Hamiltonian. This analysis is left for future work that could be an extension for this project, and is very important to see the advantage of using symmetry broken states over RHF states in simulating strongly correlated systems.

Chapter 5

Conclusion and future direction

5.1 Conclusion

This thesis has focused on speeding up the convergence time for Quantum Imaginary Time Evolution towards the ground state. While the conventional approach uses Restricted Hartree–Fock wave function as initial wave function for starting the time evolution, we initiate the time evolution with symmetry-broken wave function. Introduction of spin penalty operator to the symmetry-broken Hamiltonian is the main contribution of this work. It enlarges the energy gap between the ground state and the first excited state, allowing faster convergence with our approach. To validate the theoretical standing, we first did direct matrix exponentiation for ITE using RHF, BS2/BS3 (double/triple bond dissociation) wave functions in the N₂ molecule. It showed that RHF is still a better option for the region of weak electron correlation. However, as the bond dissociation goes on increasing, BS2 wave function is a better choice and upon further dissociation quantified by the diradical character, BS3 is the best choice. Similar behavior was observed in unitary approximated QITE with numerical simulation, gate-based implementation and implementation on a quantum processor for H₂ molecule. We also presented a method to optimize the tomography process as it serves as a bottleneck for simulating larger systems with QITE, and speed up the computational simulation speed for P4 cluster.

This work serves a good method for simulating strongly correlated systems. While molecules with qubit size larger than 6 have not been simulated with QITE (except N₂ simulation with ITE using 12 qubits), we expect it to perform better with those systems too in the region of strong correlation. In addition, for some molecules, broken symmetry wave functions provide larger initial fidelity in the region of strong correlation. This provides advantage of using symmetry-broken states in application other than QITE where the initial fidelity of the state affects the rest of the routine.

5.2 Future works

Simulating larger strongly correlated molecules to obtain their ground state remains future work beyond this thesis. As discussed earlier, larger systems require deep circuits and computationally heavy tomography process needing very large measurement circuits with many shots of measurement.

Techniques including single-shot measurement on all qubits [53], [54], [55] could be beneficial and might drastically speed-up the simulation while we use them as a substitute for full state tomography as we have to obtain expectations for many observables together. However, this is not necessarily guaranteed, as the LCU 1-norm of the terms in the Hamiltonian might be very large for larger system due to symmetry breaking and addition of penalty term. In addition, finding the expectations in QITE is not limited to single observables, but also the correlation terms, exist where two different Pauli operators come together. If that holds, techniques like classical shadow tomography [56] may not be the best choice and further investigation might be necessary.

Further, for a nearest neighbor local Hamiltonian on a cubic lattice with dimension d , the domain size D is bounded by $\mathcal{O}(C^d)$, where C is the maximum number of qubits with finite correlation. Thus, considering a molecular Hamiltonian without locality, we take domain size and maximum number of qubits with finite correlation as the total number of qubits in the Hamiltonian for exact imaginary time evolution. The number of measurements and storage needed at a time stamp is bounded by $e^{\mathcal{O}(C^d)}$ since each unitary acts on $\mathcal{O}(C^d)$ qubits. Hence, the space and time complexity increases with $e^{\mathcal{O}(C^d)}$. However, if we localize all molecular orbitals within the active space using conventional orbital localization methods [57], [58] and adopt Hamiltonian term truncation based on qubit operator locality [59], a smaller domain size D_{BS} for BS inspired QITE is expected to perform better than HF inspired QITE with domain size D_{HF} such that $D_{\text{BS}} \leq D_{\text{HF}} < C$. This has not been examined in the current scope and is set for future works.

5.3 Code Availability

The source code including data used to generate results and plots in this work can be accessed at <https://github.com/lmpawan10/QITE>.

Bibliography

- [1] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, “Simulated quantum computation of molecular energies,” *Science*, vol. 309, no. 5741, pp. 1704–1707, 2005, doi:10.1126/science.1113479.
- [2] L. C. Sala, P. A.-A. L. D. I. S.L., and M. M. Lab, “Efficient protein ground state energy computation via fragmentation and reassembly,” 2025, doi:10.48550/arXiv.2501.03766.
- [3] I. H. Kim, Y.-H. Liu, S. Pallister, W. Pol, S. Roberts, and E. Lee, “Fault-tolerant resource estimate for quantum chemical simulations: Case study on Li-ion battery electrolyte molecules,” *Physical Review Research*, vol. 4, p. 023019, 2022, doi:10.1103/PhysRevResearch.4.023019.
- [4] A. Y. Kitaev, “Quantum measurements and the abelian stabilizer problem,” 1995, doi:10.48550/arXiv.quant-ph/9511026.
- [5] D. S. Abrams and S. Lloyd, “Quantum algorithm providing exponential speed increase for finding eigenvalues and eigenvectors,” *Physical Review Letters*, vol. 83, pp. 5162–5165, 1999, doi:10.1103/PhysRevLett.83.5162.
- [6] N. E. Belaloui, A. Tounsi, R. A. Khamadja, M. M. Louamri, A. Benslama, D. E. B. Neira, and M. T. Rouabah, “Ground-state energy estimation on current quantum hardware through the variational quantum eigensolver: A practical study,” *Jurnal of Chemical Theory and Computation*, 2025, published online. doi:10.1021/acs.jctc.4c01657.
- [7] W. J. Huggins, B. A. O’Gorman, N. C. Rubin, D. R. Reichman, R. Babbush, and J. Lee, “Unbiasing fermionic quantum Monte Carlo with a quantum computer,” *Nature*, vol. 603, no. 7901, pp. 416–420, 2022, doi:10.1038/s41586-021-04351-z.
- [8] M. Motta, C. Sun, A. T. K. Tan, M. J. O’Rourke, E. Ye, A. J. Minnich, F. G. S. L. Brandão, and G. K.-L. Chan, “Determining eigenstates and thermal states on a quantum computer using quantum imaginary time evolution,” *Nature Physics*, vol. 16, pp. 205–210, 2020, doi:10.1038/s41567-020-0798-8.
- [9] K. Sugisaki, S. Yamamoto, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, “Quantum chemistry on quantum computers: A polynomial-time quantum algorithm

- for constructing the wave functions of open-shell molecules,” *The Journal of Physical Chemistry A*, vol. 120, pp. 6459–6466, 2016, doi:10.1021/acs.jpca.6b04932.
- [10] K. Sugisaki, S. Nakazawa, K. Toyota, K. Sato, D. Shiomi, and T. Takui, “Quantum chemistry on quantum computers: A method for preparation of multi-configurational wave functions on quantum computers without performing post-Hartree-Fock calculations,” *ACS Central Science*, vol. 5, pp. 167–175, 2019, doi:10.1021/acscentsci.8b00788.
 - [11] S. Lee *et al.*, “Evaluating the evidence for exponential quantum advantage in ground-state quantum chemistry,” *Nature Communications*, vol. 14, no. 1, p. 1952, 2023, doi:10.1038/s41467-023-37587-6.
 - [12] P. S. Poudel, K. Sugisaki, M. Hajdušek, and R. Van Meter, “Determining molecular ground state with quantum imaginary time evolution using broken-symmetry wave function,” in *2025 IEEE International Conference on Quantum Computing and Engineering (QCE)*, vol. 01, 2025, pp. 2279–2287, doi:10.1109/QCE65121.2025.00248.
 - [13] J. Paldus, P. Piecuch, L. Pylypow, and B. Jeziorski, “Application of Hilbert-space coupled-cluster theory to simple (H_2)₂ model systems: Planar models,” *Physical Review A*, vol. 47, no. 4, pp. 2738–2782, April 1993, doi:10.1103/PhysRevA.47.2738.
 - [14] A. Javadi-Abhari *et al.*, “Quantum computing with Qiskit,” 2024, doi:10.48550/arXiv.2405.08810.
 - [15] B. Schumacher, “Quantum coding,” *Phys. Rev. A*, vol. 51, pp. 2738–2747, Apr 1995, doi:10.1103/PhysRevA.51.2738.
 - [16] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
 - [17] W. Gerlach and O. Stern, “Der experimentelle nachweis der richtungsquantelung im magnetfeld,” *Z. Physik*, vol. 9, pp. 349–352, 1922, doi:10.1007/BF01326983.
 - [18] W. Pauli, “Zur quantenmechanik des magnetischen elektrons,” *Zeitschrift für Physik*, vol. 43, pp. 601–623, 1927.
 - [19] P. A. M. Dirac and R. H. Fowler, “The quantum theory of the electron,” *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 117, no. 778, pp. 610–624, 1928, doi:10.1098/rspa.1928.0023.
 - [20] A. Huang, “Beyond Born–Oppenheimer approximation quantum chemistry computation on a NISQ device,” *Master’s Thesis*, 2023.
 - [21] J. T. Seeley, M. J. Richard, and P. J. Love, “The Bravyi–Kitaev transformation for quantum computation of electronic structure,” *The Journal of Chemical Physics*, vol. 137, no. 22, Dec. 2012, doi:10.1063/1.4768229.

- [22] S. B. Bravyi and A. Y. Kitaev, “Fermionic quantum computation,” *Annals of Physics*, vol. 298, no. 1, p. 210–226, May 2002, doi:10.1006/aphy.2002.6254.
- [23] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, “Simulation of electronic structure hamiltonians using quantum computers,” *Molecular Physics*, vol. 109, no. 5, p. 735–750, Mar. 2011, doi:10.1080/00268976.2011.552441.
- [24] N. Moll, A. Fuhrer, P. Staar, and I. Tavernelli, “Optimizing qubit resources for quantum chemistry simulations in second quantization on a quantum computer,” *Journal of Physics A: Mathematical and Theoretical*, vol. 49, no. 29, p. 295301, Jun. 2016, doi:10.1088/1751-8113/49/29/295301.
- [25] S. Bravyi, J. M. Gambetta, A. Mezzacapo, and K. Temme, “Tapering off qubits to simulate fermionic Hamiltonians,” arXiv:1701.08213, doi:10.48550/arXiv.1701.08213.
- [26] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory.*, 1st ed. Dover Publications, Inc., 1996.
- [27] E. Coronado, “Molecular magnetism: from chemical design to spin control in molecules, materials and devices,” *Nature Reviews Materials*, vol. 5, pp. 87–104, 2020, doi:10.1038/s41578-019-0146-8.
- [28] M. Nakano and B. Champagne, “Theoretical design of open-shell singlet molecular systems for nonlinear optics,” *The Journal of Physical Chemistry Letters*, vol. 6, no. 16, pp. 3236–3256, 2015, doi:10.1021/acs.jpclett.5b00956.
- [29] Y. Umena, K. Kawakami, J.-R. Shen, and N. Kamiya, “Crystal structure of oxygen-evolving photosystem II at a resolution of 1.9 Å,” *Nature*, vol. 473, no. 7345, pp. 55–60, 2011, doi:10.1038/nature09913.
- [30] T. Spatzal, J. Schlesier, E.-M. Burger, D. Sippel, L. Zhang, S. L. A. Andrade, D. C. Rees, and O. Einsle, “Nitrogenase FeMoco investigated by spatially resolved anomalous dispersion refinement,” *Nature Communications*, vol. 7, no. 1, p. 10902, 2016, doi:10.1038/ncomms10902.
- [31] S. Lloyd, “Universal quantum simulators,” *Science*, vol. 273, pp. 1073–1078, 1996, doi:10.1126/science.273.5278.1073.
- [32] E. Campbell, “Random compiler for fast Hamiltonian simulation,” *Phys. Rev. Lett.*, vol. 123, p. 070503, Aug 2019, doi:10.1103/PhysRevLett.123.070503.
- [33] A. M. Childs and N. Wiebe, “Hamiltonian simulation using linear combinations of unitary operations,” *Quantum Info. Comput.*, vol. 12, no. 11–12, p. 901–924, nov 2012, doi:10.26421/qic12.11-12.
- [34] G. H. Low and I. L. Chuang, “Hamiltonian simulation by qubitization,” *Quantum*, vol. 3, p. 163, Jul. 2019, doi:10.22331/q-2019-07-12-163.

- [35] J. Lee, D. W. Berry, C. Gidney, W. J. Huggins, J. R. McClean, N. Wiebe, and R. Babbush, “Even more efficient quantum computations of chemistry through tensor hypercontraction,” *PRX Quantum*, vol. 2, no. 3, jul 2021, doi:10.1103/prxquantum.2.030305.
- [36] G. H. Low, R. King, D. W. Berry, Q. Han, A. E. DePrince, A. F. White, R. Babbush, R. D. Somma, and N. C. Rubin, “Fast quantum simulation of electronic structure by spectral amplification,” *Physical Review X*, vol. 15, no. 4, oct 2025, doi:10.1103/pb2g-j9cw.
- [37] R. King, G. H. Low, R. Babbush, R. D. Somma, and N. C. Rubin, “Quantum simulation with sum-of-squares spectral amplification,” 2025, doi:10.48550/arXiv.2505.01528.
- [38] M. Reiher, N. Wiebe, K. M. Svore, D. Wecker, and M. Troyer, “Elucidating reaction mechanisms on quantum computers,” *Proceedings of the National Academy of Sciences*, vol. 114, no. 29, p. 7555–7560, Jul. 2017, doi:10.1073/pnas.1619152114.
- [39] Z. Li, J. Li, N. S. Dattani, C. J. Umrigar, and G. K.-L. Chan, “The electronic complexity of the ground-state of the FeMo cofactor of nitrogenase as relevant to quantum simulations,” *The Journal of Chemical Physics*, vol. 150, no. 2, Jan. 2019, doi:10.1063/1.5063376.
- [40] J. Preskill, “Quantum computing in the NISQ era and beyond,” *Quantum*, vol. 2, p. 79, Aug. 2018, doi:10.22331/q-2018-08-06-79.
- [41] A. Anand, P. Schleich, S. Alperin-Lea, P. W. K. Jensen, S. Sim, M. Díaz-Tinoco, J. S. Kottmann, M. Degroote, A. F. Izmaylov, and A. Aspuru-Guzik, “A quantum computing view on unitary coupled cluster theory,” *Chemical Society Reviews*, vol. 51, no. 5, pp. 1659–1684, 2022, doi:10.1039/D1CS00932J.
- [42] K. Sugisaki, T. Kato, Y. Minato, K. Okuwaki, and Y. Mochizuki, “Variational quantum eigensolver simulations with the multireference unitary coupled cluster ansatz: a case study of the C_{2v} quasi-reaction pathway of beryllium insertion into a H_2 molecule,” *Physical Chemistry Chemical Physics*, vol. 24, pp. 8439–8452, 2022, doi:10.1039/D1CP04318H.
- [43] V. Kremenetski, C. Mejuto-Zaera, S. J. Cotton, and N. M. Tubman, “Simulation of adiabatic quantum computing for molecular ground states,” *The Journal of Chemical Physics*, vol. 155, no. 23, p. 234106, 12 2021, doi:10.1063/5.0060124.
- [44] R. Pauncz, *The Construction of Spin Eigenfunctions*. Kluwer Academic/Plenum Publishers, 2000.
- [45] P. Flükiger, H. P. Lüthi, S. Portmann, and J. Weber, “MOLEKEL 4.3: An interactive molecular graphics tool,” Manno, Switzerland, 2002, software package.

- [46] K. Hejazi, M. Motta, and G. K.-L. Chan, “Adiabatic quantum imaginary time evolution,” *Phys. Rev. Res.*, vol. 6, p. 033084, Jul 2024, doi:10.1103/PhysRevResearch.6.033084.
- [47] A. W. Harrow, A. Hassidim, and S. Lloyd, “Quantum algorithm for linear systems of equations,” *Phys. Rev. Lett.*, vol. 103, p. 150502, Oct 2009, doi:10.1103/PhysRevLett.103.150502.
- [48] F. Zahariev *et al.*, “The general atomic and molecular electronic structure system (GAMESS): Novel methods on novel architectures,” *Journal of Chemical Theory and Computation*, vol. 19, no. 20, pp. 7031–7055, October 2023, doi:10.1021/acs.jctc.3c00379.
- [49] J. R. McClean *et al.*, “OpenFermion: The electronic structure package for quantum computers,” *Quantum Science and Technology*, vol. 5, no. 3, p. 034014, June 2020, doi:10.1088/2058-9565/ab8ebc.
- [50] D. Doehnert and J. Koutecky, “Occupation numbers of natural orbitals as a criterion for biradical character. different kinds of biradicals,” *Journal of the American Chemical Society*, vol. 102, no. 6, pp. 1789–1796, 1980, doi:10.1021/ja00526a005.
- [51] G. Li, Y. Ding, and Y. Xie, “Tackling the qubit mapping problem for NISQ-era quantum devices,” in *Proceedings of the Twenty-Fourth International Conference on Architectural Support for Programming Languages and Operating Systems*, ser. ASPLOS ’19. New York, NY, USA: Association for Computing Machinery, 2019, p. 1001–1014, doi:10.1145/3297858.3304023.
- [52] “Compute resources, IBM quantum,” https://quantum.cloud.ibm.com/computers?system=ibm_kingston, accessed: 2025-10-18.
- [53] T. J. Evans, R. Harper, and S. T. Flammia, “Scalable bayesian Hamiltonian learning,” 2019, doi:10.48550/arXiv.1912.07636.
- [54] H.-Y. Huang, R. Kueng, and J. Preskill, “Information-theoretic bounds on quantum advantage in machine learning,” *Phys. Rev. Lett.*, vol. 126, p. 190505, May 2021, doi:10.1103/PhysRevLett.126.190505.
- [55] A. Montanaro, “Learning stabilizer states by Bell sampling,” 2017, doi:10.48550/arXiv.1707.04012.
- [56] H.-Y. Huang, R. Kueng, and J. Preskill, “Predicting many properties of a quantum system from very few measurements,” *Nature Physics*, vol. 16, no. 10, p. 1050–1057, Jun. 2020, doi:10.1038/s41567-020-0932-7.
- [57] S. F. Boys, “Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another,” *Reviews of Modern Physics*, vol. 32, pp. 296–299, 1960, doi:10.1103/RevModPhys.32.296.

- [58] J. Pipek and P. G. Mezey, “A fast intrinsic localization procedure applicable for ab initio and semiempirical linear combination of atomic orbital wave functions,” *The Journal of Chemical Physics*, vol. 90, no. 9, pp. 4916–4926, May 1989, doi:10.1063/1.456588.
- [59] K. Sugisaki, S. Kanno, T. Itoko, R. Sakuma, and N. Yamamoto, “Hamiltonian simulation-based quantum-selected configuration interaction for large-scale electronic structure calculations with a quantum computer,” 2024, doi:10.48550/arXiv.2412.07218.