Born-Oppenheimer Molecular Dynamics using the Variational Quantum Eigensolver

Nicolas Poirier

Department of Mechanical Engineering

McGill University

Montreal, QC

December 2019

A thesis submitted in partial fulfillment of the requirements of the degree of Master of Engineering

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

Superscripts and Accents

$\ddot{(\cdot)}$	Second-order time derivative
$\dot{(\cdot)}$	First-order time derivative
$\hat{(\cdot)}$	Operator (Note that not all operators have this accent.)
$\langle\cdot\rangle$	Expectation value
$\langle\cdot\vert$	Bra vector
$(\cdot)^{AO}$	Atomic orbital basis
$(\cdot)^{MO}$	Molecular orbital basis
$(\cdot)^{MSO}$	Molecular spin-orbital basis
$\vert\cdot\rangle$	Ket vector
$\widehat{(\cdot)}$	Estimator

Abbreviations

AIMD	*Ab initio* molecular dynamics
AO	Atomic orbital
BOMD	Born-Oppenheimer molecular dynamics
CGTO	Contracted Gaussian-type orbital
DIIS	Direct inversion in the iterative subspace
FCI	Full configuration interaction
MCSCF	Multiconfigurational self-consistent field

MO	Molecular orbital
MO	Molecular spin-orbital
MP2	Second-order Møller-Plesset perturbation theory
PEA	Phase estimation algorithm
PES	Potential energy surface
QASM	Quantum Assembly Language
QFT	Quantum Fourier transform
RHF	Restricted Hartree-Fock
SPSA	Simultaneous perturbation stochastic approximation
TDSE	Time-dependent Schrödinger equation
TISE	Time-independent Schrödinger equation
UCC	Unitary coupled cluster
UCCSD	Unitary coupled cluster singles-and-doubles
VQE	Variational quantum eigensolver

Abstract

Using the laws of quantum mechanics, it is possible to simulate the time evolution of molecular systems. The results of these simulations are of great importance to fields such as combustion, materials science, and drug design. They can be used to determine chemical properties such as critical points and reaction rates. However, the dimension of the Hilbert space that describes the wave function of a molecule scales exponentially with the size of the system. A solution to this scaling problem is to use quantum computers to simulate molecules. In this thesis, a strategy for conducting molecular dynamics simulations is explained in detail. This strategy involves using a hybrid algorithm known as the variational quantum eigensolver (VQE). The algorithm takes advantage of the computational power of both classical and quantum computers. It works by varying optimization parameters associated with the electronic wave function until the total electronic energy is minimized. Born-Oppenheimer molecular dynamics is then employed to determine the motion of the nuclei based on the results of the VQE. This method was tested by simulating the time evolution of H_2 using a quantum computer simulator. Attempts were also made to repeat this simulation with a superconducting quantum computer.

Additionally, techniques to improve the efficiency of the VQE and the BOMD method are presented in this thesis. The first strategy consists in varying the number of quantum measurements during the VQE's optimization procedure. It was demonstrated that the number of measurements needed could be reduced by a factor of almost 3 for H_2 by varying the number of measurements based on an exponential relationship. It was also shown that the initial guess for the optimization parameters could be improved during a BOMD simulation

by utilizing results from previous time points. The average relative difference between the guessed parameters and the optimal ones was 2.5% when using a linear extrapolation technique to simulate the time evolution of H_2. On the other hand, it was 35% when employing perturbation theory to generate guesses.

Abrégé

À l'aide des lois de la mécanique quantique, il est possible de simuler l'évolution temporelle de systèmes moléculaires. Les résultats de ces simulations sont d'une grande importance dans plusieurs domaines tels que la combustion, la science des matériaux et la conception de médicaments. Ils peuvent être utilisés afin de déterminer des propriétés chimiques comme un point critique et la vitesse de réaction. Toutefois, la dimension de l'espace de Hilbert décrivant la fonction d'onde s'accroît exponentiellement avec la taille du système. Une solution à ce problème est d'utiliser un ordinateur quantique au lieu d'un ordinateur classique pour simuler des molécules. Dans cette thèse, une stratégie pour simuler la dynamique des molécules sera expliquée en détail. Cette stratégie implique l'utilisation d'un algorithme hybride connu sous le nom de «Variational Quantum Eigensolver» (VQE). Ce dernier exploite la puissance non seulement d'un ordinateur quantique, mais aussi d'un ordinateur classique. Il fonctionne en variant les paramètres d'optimisation associés à une fonction d'onde électronique jusqu'à ce que l'énergie électronique totale soit minimisée. La dynamique moléculaire de Born-Oppenheimer (BOMD) est ensuite employée pour déterminer la position des noyaux d'après les résultats du VQE. Cette méthode fut testée en simulant l'évolution temporelle d'une molécule de H_2 à l'aide d'un simulateur d'ordinateur quantique. Ce test fut aussi tenté avec un ordinateur quantique composé de circuits supraconducteurs.

Additionnellement, des techniques pour améliorer l'efficacité du VQE et de la BOMD sont aussi présentées dans cette thèse. La première stratégie consiste à varier le nombre de mesures quantiques lors de la procédure d'optimisation du VQE. Il fut démontré que le nombre requis de mesures peut être réduit par un facteur de presque 3, pour une molécule

d'hydrogène, en variant le nombre de mesures d'après une fonction exponentielle. Il fut aussi illustré que la valeur de départ des paramètres d'optimisation pouvait être amélioré pendant la BOMD en utilisant les résultats des temps précédents. Par exemple, la différence relative moyenne entre les paramètres de départ et les paramètres optimaux pour la simulation d'une molécule d'hydrogène fut de 2,5% en employant une technique d'extrapolation linéaire. En terme de comparaison, elle fut de 35% pour des paramètres de départ déterminés à l'aide de la théorie de la perturbation.

Acknowledgments

Without the help and guidance from several individuals, the completion of the work presented in this thesis would not have been possible. I would first like to thank my supervisor Prof. Luc Mongeau for providing the guidance and the support necessary for undertaking this research project. I am also immensely thankful to my co-supervisor Dr. Alireza Najafi-Yazdi who provided me with all the tools I needed to achieve the daunting task of using quantum computing to simulate molecules. Not only was he a great teacher for topics ranging from quantum mechanics to numerical methods, but he also encouraged me to challenge myself and aim for excellence. Furthermore, his teaching of clean and professional programming strategies were vital to the completion of this project. I would also like to thank him for his financial support.

Additionally, I am grateful to Prof. Rustam Khaliullin with who I had insightful discussions about quantum mechanics and computational chemistry. Special thanks also go to Dr. Kevin Lalumière for helping me understand several aspects of quantum computing and for providing me with tips for coding more efficiently. Moreover, I would like to thank the team at Anyon Systems for taking their time to explain various components of a superconducting quantum computer, ranging from the cryogenics system to the electronics.

Lastly, I am thankful to McGill University for offering me the Graduate Excellence Fellowship. I am also grateful to IBM for making their quantum computers accessible to the public, free of charge.

Contributions

The author of this thesis contributed to the original work which is summarized in the following list:

- A strategy for reducing the number of quantum measurements in the VQE's optimization procedure by varying the uncertainty associated with the directional derivative of energy.
- A simulation of the time evolution of H_2 using BOMD and the VQE on a quantum computer simulator.
- A comparison of the methods used to select the initial guess for the VQE's optimization parameters in the BOMD approach.
- An ansatz with a short circuit depth for entangling the reference state of H_2 in the STO-3G basis set.

CHAPTER 1
Introduction

1.1 Motivation

Quantum mechanics is needed to predict the dynamics of atoms and molecules at small scales [1–3]. One of the postulates of quantum mechanics states that any system of particles can be described using a wave function Ψ [4]. Knowing the wave function, which is also called the quantum state, it is possible to determine any property of the system.

To determine Ψ and how it evolves in time t, the time-dependent Schrödinger equation (TDSE) can be employed. This equation is defined as [1, 4]:

$$i\hbar \frac{\partial \Psi}{\partial t} = \hat{H}\Psi, \tag{1.1}$$

where $\hbar$ is the reduced Planck constant and $\hat{H}$ is the Hamiltonian of the system. The Hamiltonian corresponds to the sum of the kinetic energy operator $\hat{T}$ and the potential energy operator $\hat{V}$.

Unfortunately, attempting to solve the TDSE using a classical computer is challenging due to the scaling of the resources required to describe a quantum state. For a single particle, the quantum state can be described using a vector $|\Psi_{sub,0}\rangle$ in a Hilbert space with dimension d_0. If a second particle with a Hilbert space of dimension d_1 is added to the system, one might expect from classical mechanics that the Hilbert space of the system would have the dimension $d_0 + d_1$. However, in quantum mechanics, a system $|\Psi\rangle$ composed of two subsystems, $|\Psi_{sub,0}\rangle$ and $|\Psi_{sub,1}\rangle$, is expressed in terms of a tensor product $\otimes$ as shown in the following equation:

$$|\Psi\rangle = |\Psi_{sub,0}\rangle \otimes |\Psi_{sub,1}\rangle. \tag{1.2}$$

The Hilbert space of such a system has dimension $d_0 d_1$. For a system composed of n_{sub} subsystems, the dimension of the Hilbert space corresponds to $d_0 d_1 \ldots d_{n_{sub}-2} d_{n_{sub}-1}$ [5]. This means that the Hilbert space scales exponentially with the number of particles in a quantum system.

To circumvent this issue, scientists have developed methods over the years to compute approximate solutions to the TDSE [6]. These methods take advantage of techniques such as the Hartree-Fock method [2, 7], perturbation theory [8, 9], or density functional theory (DFT) [10, 11]. While they provide sufficiently accurate results in some instances, they fail in a number of other cases. For example, the first two approaches are typically unable to correctly predict dissociation energies [12] while the last one tends to fail for transition states and covalent bonding [13].

As an alternate solution to this dimensionality problem, Richard Feynman proposed using computers that take advantage of the laws of quantum mechanics [14] in order to simulate quantum systems. On such a computer, the amount of data that can be processed scales exponentially with the number of quantum bits (qubits) available. Each of these qubits corresponds to a two-level system. The state of a single qubit is the following [15]:

$$|\Psi\rangle = \gamma_0|0\rangle + \gamma_1|1\rangle, \tag{1.3}$$

where $|\Psi\rangle$ is normalized and γ is a complex coefficient. For a system of two qubits, the quantum state becomes the tensor product of two single-qubit states. The resulting state can be represented as:

$$|\psi\rangle = \gamma_{00}|00\rangle + \gamma_{01}|01\rangle + \gamma_{10}|10\rangle + \gamma_{11}|11\rangle. \tag{1.4}$$

As more qubits are added, $|\psi\rangle$ becomes the tensor product of a larger number of single-qubit states. This means that a system containing N_q qubits can store 2^{N_q} values [15, 16].

Constructing a quantum computer that can run quantum algorithms is proving to be a challenging task, however. This is partly because quantum computers are not closed systems. In reality, they interact with their surrounding environment. Some of these interactions are necessary, like when measuring the state of a qubit. However, other interactions are undesired and can adversely alter a qubit's quantum state. Those unwanted changes are known as quantum noise or decoherence [15]. To quantify the impact of decoherence, characteristic times known as decoherence times τ_Q can be used to quantify the rapidity at which the environment affects the state of a quantum computer. The decoherence times therefore impose a limit on the maximum number of quantum operations that can be applied to a qubit [17]. To be more precise, the maximum number of operations that can be applied is proportional to τ_Q/τ_O, where τ_O corresponds to the time required to perform an operation [15]. Fortunately, there exist methods for correcting errors such as those which are due to decoherence. One approach, known as quantum error correction, involves employing multiple qubits to encode a single-qubit state. Using this strategy, it is then possible to measure errors and apply corrections without destroying the encoded state [18]. According to an estimate made by Fowler *et al.* [19], though, the number of qubits needed to encode a single-qubit state is on the order of 10^3 to 10^4. In contrast, today's largest quantum computers only contain a few dozen qubits [20].

Despite these challenges, simulations of quantum systems using multiple qubits have been successfully carried out using trapped ions [21] and superconducting circuits [22]. However, no one has yet been able to simulate the time evolution of a molecule using a quantum computer. Successfully completing such a computation would be an important milestone in the quest to solve quantum chemistry problems that are intractable on classical computers. Moreover, it is worth noting that these problems have a wide range of applications. For example, it would become possible to simulate combustion reactions more accurately. Such simulations would help provide a better understanding of combustion reaction mechanisms and they could be used to determine reaction rates [23]. This would provide additional tools to engineers who are designing increasingly efficient combustion engines.

Another area that would greatly benefit from the ability to solve complex molecular dynamics problems is materials science. For example, it would become easier to study materials under extreme conditions such as high temperatures, high pressures, and radiation. It can be very difficult or even impossible to study materials which are under such conditions in a laboratory setting. These conditions are encountered when trying to generate energy using nuclear fusion for instance [24]. A key concern in nuclear fusion power plants is the ability of the first wall to resist high levels of neutron radiation. This radiation can lead to a reduction of the wall's fracture toughness. Being able to efficiently simulate the effects of neutron radiation on the wall would help select the optimal wall material [25]. In a similar manner, the computational power of quantum computers could be harnessed to help study oxidation reactions such as corrosion [26]. Additionally, it could be used to develop new gate dielectrics for field effect transistors [27] or to find new materials that could improve the efficiency of solar cells [28].

Lastly, the arrival of quantum computers could also have a great impact on the field of medicine. For example, predicting how proteins fold could be done more accurately. Being able to make those predicitons is an invaluable tool when designing new drugs [29]. Furthermore, quantum computers could help design artificial molecular machines. These machines could be utilized to replicate the behaviour of natural biomolecular machines or, even, to interact with them [30]. It has already been shown that artificial molecular machines can be used to alter the mechanical properties of materials [31]. They can also act as photoswitchable catalysts, which only become active when exposed to light [32].

1.2 Literature Review: *Ab Initio* Molecular Dynamics

The focus of this thesis will be on *ab initio* molecular dynamics (AIMD) methods. The term *ab initio* indicates that the methods rely on first principles, which in this case are the laws of quantum mechanics [6]. At the other end of the spectrum, there also exist a wide range of molecular dynamics approaches which use parameterized potentials to predict the time evolution of molecules. Although they are less computationally expensive than AIMD methods, they rely on a set of parameters which have been fitted to experimental data. As a consequence, they tend to provide inaccurate results when used to simulate systems that are markedly different from the experiments used to determine the parameters [33, 34].

One of the first techniques for simulating the time evolution of molecules was developed by Ehrenfest [35] in 1927. It works by treating all the nuclei in the system as classical point particles and removing the kinetic energy operator for the nuclei in the Hamiltonian. In this form, the Hamiltonian is known as the electronic Hamiltonian $\hat{H}_e$. To predict the position of the nuclei at future points in time, the TDSE and the electronic Hamiltonian are used to determine a time-dependent wave function $\Psi_e(t)$. This wave function is then employed to determine the force acting on each nuclei. Afterwards, the positions of the nuclei at the next time point are found by integrating Newtonian equations of motion. The advantage of this method is that the Hamiltonian only needs to be diagonalized at the beginning of the simulation. However, the time steps must be chosen based on the motion of the electrons which evolve at a much shorter time scale than the nuclei [6].

An alternative approach is to replace the TDSE in Ehrenfest molecular dynamics by the time-independent Schrödinger equation (TISE), which corresponds to:

$$\hat{H}_e\Psi_e = E\Psi_e, \tag{1.5}$$

where E is called the total electronic energy. This different method is known as Born-Oppenheimer molecular dynamics (BOMD) [6]. In this case, since the electronic wave function is solved using the TISE, the time step is selected based on the time scale associated with the motion of the nuclei. The drawback though is that the electronic Hamiltonian must be diagonalized at each time step. This is a computationally expensive task for a classical computer.

In 1985, Car and Parrinello [36] developed an approach that can be considered a compromise between Ehrenfest molecular dynamics and BOMD. This method works by adding a fictitious kinetic energy term, which is associated with the dynamics of the electronic wave function, to the Lagrangian describing the system. The resulting equations of motion then describe not only the dynamics of the nuclei with mass M_I, but also the dynamics of components of Ψ_e with fictitious mass μ. The consequence is that the electronic Hamiltonian only needs to be diagonalized once like for Ehrenfest molecular dynamics, but a longer time step can be chosen. This time step must still be about an order of magnitude smaller than one that is selected for the BOMD method however [6].

As explained in the previous section, the issue when using one of these three methods on a classical computer is that the computational resource requirements scale exponentially with the size of the simulated system. As an altenative, Lloyd [37] showed in 1996 that quantum computers could be used to simulate the evolution of a quantum system in polynomial time. Mathematically, this time evolution can be described as:

$$|\Psi\left(t\right)\rangle = \hat{U}\left(t\right)|\Psi\left(0\right)\rangle = \exp\left(-i\hat{H}t\right)|\Psi\left(0\right)\rangle, \tag{1.6}$$

where $\hat{U}\left(t\right)$ is a unitary operator. In his seminal paper, Lloyd demonstrated that this operator can be approximated using Trotter decomposition, as shown in the following equation:

$$\hat{U}\left(t\right) \approx \left(\exp\left(-\frac{i\hat{H}_0 t}{j}\right)\ldots\exp\left(-\frac{i\hat{H}_{j-1}t}{j}\right)\right)^j. \tag{1.7}$$

The operator $\hat{U}\left(t\right)$ can be expressed in this manner as long as the Hamiltonian can be written as the following sum:

$$\hat{H} = \sum_{i=0}^{j-1}\hat{H}_i, \tag{1.8}$$

where each term $\hat{H}_i$ only involves a portion of the qubits on a quantum computer. To be able to simulate the time evolution of $|\Psi\left(t\right)\rangle$ in polynomial time, the number of terms in the Hamiltonian must also scale polynomially with the size of the system.

Following this breakthrough, Wiesner [38] and Zalka [39] suggested representing a system's wave function using 2^{N_q} points in space. These points can then be stored in the quantum state of N_q qubits. Furthermore, since the kinetic portion of the Hamiltonian is diagonal in the momentum basis, they suggested using the quantum Fourier transform

(QFT) [40] when applying $\hat{U}(t)$. The purpose of the transform and its inverse is to change the basis of the qubits from the position basis to the momentum basis and vice versa. With this approach, Lidar and Wang [41] showed in 1999 that it was possible to compute the thermal rate constants of chemical reactions in polynomial time. Then, in 2008, Kassal *et al.* [42] improved Zalka and Wiesner's method and demonstrated that molecular systems could be simulated in $O\left(N_{tot}^2 \varepsilon_V^2\right)$ time, where N_{tot} is the number of particles and ε_V is the precision of the discretized potential energy function. They also developed strategies for efficiently determining reaction probabilities and state-to-state transition probabilities.

The previous approaches for simulating the time evolution of a quantum state all use the first quantization language. In this formulation of quantum mechanics, the wave function can be expressed in terms of position or momentum coordinates. An alternate strategy, known as second quantization, is to express $|\Psi\rangle$ as an occupation-number vector. This vector is shown in the following equation:

$$|\Psi\rangle = |k_0 k_1 \ldots k_{N_{SO}-1}\rangle, \tag{1.9}$$

where N_{SO} is the number of spin-orbitals that could be used to represent the wave function. If k_i is 1, it means that the ith spin-orbital is occupied, while if it is 0, it is unoccupied [12]. Using the second quantization language, Abrams and Lloyd [43], developed an algorithm in 1999 to simulate the time evolution of a quantum state on a quantum computer. In their algorithm, the state of the qubits is set such that it represents the occupation-number vector associated with a quantum state. The simulation then proceeds by applying the unitary operator given in eq. (1.7). Like for the first quantization methods, their algorithm can complete a simulation in polynomial time. Since then, numerous improvements have been made to improve the scaling of this approach, which is now known as the phase estimation algorithm (PEA) [44–46].

Another strategy for solving the time-dependent Schrödinger equation in polynomial time is to combine the power of classical computers and quantum computers. In 2017, Li and Benjamin [47] developed an algorithm that employs a variational approach for solving the TDSE. It works by expressing the quantum state in terms of parameters $\boldsymbol{\theta}$. The evolution of these parameters is then described using the following Euler-Lagrange equation:

$$\sum_q M_{k,q}\dot{\theta}_q = P_k. \tag{1.10}$$

Here, $\dot{\theta}_q$ is the time derivative of θ_q, while $M_{k,q}$ and P_k are real coefficients which are defined as:

$$M_{k,q} = i\frac{\partial\langle\Psi|}{\partial\theta_k}\frac{\partial|\Psi\rangle}{\partial\theta_q} + \text{H.c.}, \tag{1.11}$$

and

$$P_k = \frac{\partial\langle\Psi|}{\partial\theta_k}\hat{H}|\Psi\rangle + \text{H.c.} \tag{1.12}$$

In this case, H.c. stands for Hermitian conjugate and $\langle\Psi|$ is the Hermitian conjugate of $|\Psi\rangle$. To begin the algorithm, the coefficients $M_{k,q}$ and P_k are evaluated using a quantum computer at the initial time point. The Euler-Lagrange equations are then solved on a classical computer in order to determine new parameters $\boldsymbol{\theta}$ at the next time point. This process is repeated until the total time is reached. At this stage, the final quantum state can be obtained using the final parameters.

Finally, it is also possible to conduct molecular dynamics simulations by using a quantum computer to solve the TISE at every time point. A classical computer computer can then be employed to solve the Newtonian equations of motion for BOMD. At the moment, there exist a few methods for solving the TISE using quantum computers. These will be introduced in the following section. For instance, it was mentioned by O'Malley *et al.* [48] that one could run molecular dynamics simulations by using the variational quantum eigensolver (VQE) to compute the forces acting on the nuclei of a molecule.

1.3 Literature Review: Electronic Structure Methods

On a classical computer, the most expensive part of the BOMD method is to diagonalize the electronic Hamiltonian in order to determine the electronic structure of the molecule. However, Aspuru-Guzik *et al.* [44] showed in 2005 that it was possible to use the phase estimation algorithm to solve the electronic structure problem in polynomial time. To determine the total electronic energy E, they took advantage of the fact that, when applied to $|\Psi_e\rangle$, the unitary operator $\hat{U}(t)$ given in eq. (1.6) can be written as:

$$\hat{U}(t)\,|\Psi_e\rangle = \exp\left(i\hat{H}_e t\right)|\Psi_e\rangle = \exp(iEt)\,|\Psi_e\rangle = \exp(i2\pi\varphi)\,|\Psi_e\rangle. \tag{1.13}$$

Here, φ corresponds to the phase of the quantum state's eigenvalue. Since then, several variants of this algorithm have been developed to increase its efficiency [49, 50]. So far, the PEA has been used to determine the dissociation curve of H_2 using photonic devices [51, 52] and superconducting qubits [48].

As an alternative, Peruzzo *et. al* [53] proposed a method known as the variational quantum eigensolver (VQE) in 2014. This algorithm takes advantage of both classical and quantum computers. It works by first parameterizing the electronic wave function. Given some parameters $\boldsymbol{\theta}$, the total electronic energy E is measured using a quantum computer. The parameters are then updated using a classical optimization algorithm. This process is repeated until E is minimized. One of the main advantages of this method is that it requires a smaller number of quantum gates than the PEA for small molecules. Additionally, its variational nature makes it more resilient to noise than the PEA [54]. In their paper introducing the VQE, Peruzzo *et al.* were able to obtain the potential energy surface (PES) of HeH^+ using a quantum photonic chip. The algorithm has also been tested on superconducting quantum computers. For example, O'Malley *et al.* [48] used it to determine the PES of H_2, while Kandala *et al.* [55] achieved the same goal for H_2, LiH, and BeH_2. On trapped-ion quantum computers, Shen *et al.* [56] obtained the dissociation curve for HeH^+, while Hempel *et. al* [57] did the same for H_2 and LiH. Total electronic energy values for H_2O were also computed by Nam *et al.* [58] using trapped-ions.

1.4 Objectives

As demonstrated in Section 1.1, quantum computers promise to solve, in polynomial time, a plethora of chemistry problems which have engineering applications but are intractable on classical computers. However, no one has yet been able to take advantage of quantum computers to simulate the time evolution of molecular systems. For that reason, the goal of this research project was to develop numerical schemes and methods that are able to perform molecular dynamics simulations using today's small and noisy quantum computers. To achieve this objective, it was decided to use the BOMD method along with a quantum computer algorithm that can find approximate solutions to the TDSE. This approach was selected since it has already been demonstrated that the PES of molecules could be determined using real quantum devices. To solve the electronic structure problem, the VQE was chosen because it is more resilient to noise than the PEA and it requires less quantum gates for small molecules. As for the second objective, it was to simulate the dynamics of H_2 using the BOMD method on a quantum computer simulator. There are a few reasons why H_2 was selected to test the quantum algorithm. Firstly, molecular hydrogen only contains two electrons. This number is large enough for H_2 to experience interelectronic interactions [12]. Yet, it is sufficiently small to allow the time evolution of H_2 to be predicted using today's quantum computers [48]. Furthermore, simulating a small number of electrons makes analyzing the results of a simulation an easier task. Other sub-objectives were also defined in order to improve the efficiency of the VQE and the BOMD method. One of these sub-objectives was to minimize the number of quantum measurements required during the optimization procedure of the VQE. Another sub-objective was to use results from previous time points in the BOMD method in order to compute more accurate initial guesses for the VQE's optimization parameters.

1.5 Organization of the Thesis

In this thesis, an overview of Born-Oppenheimer molecular dynamics is first presented in Chapter 2. Afterwards, a detailed description of the variational quantum eigensolver is given in Chapter 3. Results of VQE simulations conducted to determine the potential energy surface of H_2 are analyzed in Chapter 4. These calculations were performed using a quantum computer simulator. A strategy for reducing the number of quantum measurements in the VQE's optimization procedure is also presented in this chapter. As for Chapter 5, it details how to compute the energy gradient and integrate the equations of motion in the BOMD method. In Chapter 6, a molecular dynamics simulation of H_2 is conducted using a quantum computer simulator. A strategy for improving the initial guess of the optimization parameters in the BOMD method is also described and tested. As for Chapter 7, it contains the results of an attempt to run a molecular dynamics simulation using a superconducting quantum computer. To complete this thesis, some concluding remarks as well as suggestions for future work are given in Chapter 8.

CHAPTER 2
Overview of Born-Oppenheimer Molecular Dynamics

2.1 Molecular Hamiltonian

As with other AIMD methods, the goal of Born-Oppenheimer molecular dynamics (BOMD) is to find an approximate solution to the TDSE. Assuming relativistic effects are negligible, the Hamiltonian $\hat{H}$ for a system of molecules is defined as [12]:

$$\hat{H} = \sum_{I=0}^{N-1} \hat{T}_{nuc,I} + \sum_{i=0}^{n-1} \hat{T}_{elec,i} + \sum_{I<J}^{N-1} \hat{V}_{nuc-nuc,I,J} + \sum_{I,i}^{N-1,n-1} \hat{V}_{nuc-elec,I,i} + \sum_{i<j}^{n-1} \hat{V}_{elec-elec,i,j}, \quad (2.1)$$

where N and n are the number of nuclei and electrons in the system, respectively. The kinetic energy operator of nucleus I corresponds to

$$\hat{T}_{nuc,I} = -\frac{\hbar^2}{2M_I}\nabla_I^2, \quad (2.2)$$

while the kinetic energy operator of electron i corresponds to

$$\hat{T}_{elec,i} = -\frac{\hbar^2}{2m_e}\nabla_i^2. \quad (2.3)$$

In the two previous equations, $\hbar$ is the reduced Planck constant, M_I is the mass of nucleus I, m_e is the mass of an electron, and ∇^2 is the Laplacian with respect to the spatial coordinates of either nucleus I or electron i. The operator for the potential energy due to Coulomb forces between nuclei I and J is defined as

$$\hat{V}_{nuc-nuc,I,J} = \frac{e^2 Z_I Z_J}{4\pi\varepsilon_0|\mathbf{R}_I - \mathbf{R}_J|} \quad , \quad (2.4)$$

where e is the elementary electric charge and ε_0 is the vacuum permittivity. Additionally, Z_I and $\mathbf{R}_I$ are the atomic number and the spatial coordinates of nucleus I, respectively. Lastly, the nuclear-electron potential operator $\hat{V}_{nuc-elec,I,i}$ and the electron-electron potential operator $\hat{V}_{elec-elec,i,j}$, which are also based on Coulomb forces, can be found using

$$\hat{V}_{nuc-elec,I,i} = -\frac{e^2 Z_I}{4\pi\varepsilon_0 |\mathbf{R}_I - \mathbf{r}_i,|}, \tag{2.5}$$

and

$$\hat{V}_{elec-elec,i,j} = \frac{e^2}{4\pi\varepsilon_0 |\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.6}$$

where $\mathbf{r}_i$ are the spatial coordinates of electron i.

2.2 Justification of Electronic and Nuclear Wave Function Decoupling

When deriving the equations for BOMD, a few assumptions need to be made. One of these assumptions is that the total wave function Ψ can be expressed in terms of electronic wave functions $\Psi_{e,l}$ and nuclear wave functions $\Psi_{n,l}$, as shown in the following equation [59]:

$$\Psi\left(\mathbf{r}, \mathbf{R}, t\right) \approx \sum_l \Psi_{e,l}\left(\mathbf{r}, \mathbf{R}\right) \Psi_{n,l}\left(\mathbf{R}, t\right). \tag{2.7}$$

It is further assumed that the electronic wave functions are solutions to the time-independent Schrödinger equation (TISE)

$$\hat{H}_e \Psi_{e,l} = E_l \Psi_{e,l}, \tag{2.8}$$

where E_l are electronic eigenenergies and $\hat{H}_e$ is the electronic Hamiltonian. This Hamiltonian is the same as the molecular one except that the kinetic energy of the nuclei is not included.

The assumption that the total wave function can be expressed in terms of solutions to the TISE is typically valid since electrons evolve on a shorter time scale than the nuclei.

This is primarily due to the relatively high kinetic energy of electrons. The result is that the electrons can adjust to the position of the nuclei rapidly. In other words, the nuclei appear to be fixed in comparison to the electrons. This means that the electronic wave functions $\Psi_{e,l}$ practically reach steady state for every new configuration of the nuclear wave functions $\Psi_{n,l}$.

The difference in magnitude between the kinetic energy of the electrons and the nuclei can be expressed quantitatively as well. The upcoming analysis follows the one that was presented by Goodisman [60] for diatomic molecules. As a starting point for estimating the electronic kinetic energy T_{elec} of a diatomic molecule such as H_2, the Heisenberg uncertainty principle [61] can be employed. For position and momentum, this principle is expressed as [62]:

$$\langle(\Delta x)^2\rangle\langle(\Delta p)^2\rangle \geq \frac{\hbar^2}{4}, \tag{2.9}$$

where $\langle(\Delta x)^2\rangle$ is the variance of position and $\langle(\Delta p)^2\rangle$ is the variance of momentum. For a diatomic molecule, the variance in the position of the electrons is approximately equal to the interatomic distance d. As for the variance of momentum, it is defined as:

$$\langle(\Delta p)^2\rangle = \langle p^2\rangle - \langle p\rangle^2, \tag{2.10}$$

where $\langle p^2\rangle$ is the expectation value of momentum squared, while $\langle p\rangle$ is the expectation value of momentum. Since electrons do not have a preferred direction in which to move, $\langle p\rangle$ is zero. This means that $\langle p^2\rangle$ is of the following order of magnitude:

$$\langle p^2\rangle \sim \left(\frac{\hbar}{d}\right)^2. \tag{2.11}$$

Furthermore, because the kinetic energy of an electron can be written in terms of momentum, the size of T_{elec} can be approximated as:

$$T_{elec} = \frac{\langle p^2 \rangle}{2m_e} \sim \frac{\hbar^2}{2m_e d^2}. \tag{2.12}$$

As for the nuclei, their kinetic energy T_{nuc} is mainly dependent on their vibrational energy E_{vib} [60]. To approximate the vibrational energy, a harmonic oscillator can be used. The energy of a harmonic oscillator is proportional to $\hbar\omega$, where the angular frequency ω corresponds to:

$$\omega = \sqrt{\frac{k}{M_I}}. \tag{2.13}$$

Here, the force constant k can be estimated by noting that stretching the molecule for a distance of approximately d will cause it to dissociate. The energy required for dissociation is nearly equal to the ionization energy of an electron, which in turn is approximately equal to the kinetic energy of an electron. Using eq. (2.12), the dissociation energy E_{dis} can be written as:

$$E_{dis} = \frac{1}{2}kd^2 \sim \frac{\hbar^2}{2m_e d^2}. \tag{2.14}$$

Solving for the force constant and substituting it into eq. (2.13) gives:

$$\omega \sim \frac{\hbar}{d^2\sqrt{M_I m_e}}. \tag{2.15}$$

This means that the vibrational energy of the nuclei has the following order of magnitude [60]:

$$E_{vib} \sim \hbar\omega \sim \frac{\hbar^2}{d^2\sqrt{M_I m_e}}. \tag{2.16}$$

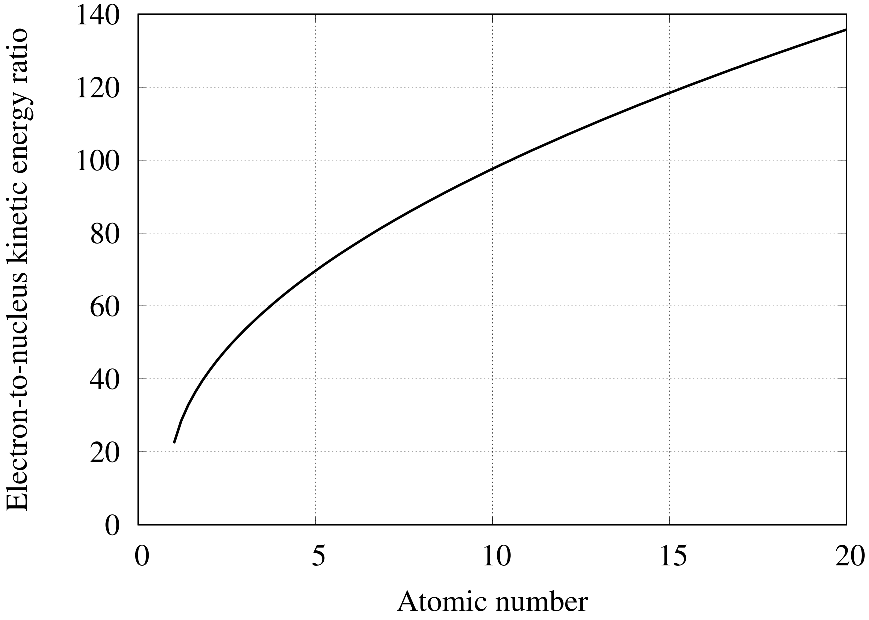

Figure 2–1: A plot showing the effects of atomic number on the ratio T_{elec}/T_{nuc}.

As a result, the ratio of the kinetic energy of an electron to the one of a nuclei is of the following order of magnitude:

$$\frac{T_{elec}}{T_{nuc}} \sim \frac{1}{2}\sqrt{\frac{M_I}{m_e}}. \tag{2.17}$$

Since a proton is nearly 2000 times the mass of an electron [63], the ratio is on the order of 20 to 1 for a hydrogen nucleus. As illustrated in Fig. 2–1, it is even larger for more massive nuclei. For instance, T_{elec}/T_{nuc} is nearly 60 to 1 for a lithium nucleus, while it is about 90 to 1 for an oxygen nucleus. At worst, T_{elec} is at least an order of magnitude larger than T_{nuc}. Therefore, it is justified to use the TISE to solve for $\Psi_{e,l}$ when simulating the molecular dynamics of molecules.

2.3 Derivation

Having justified the use of the electronic TISE, it is now possible to substitute the total wave function expansion, given in eq. (2.7), into the TDSE, shown in eq. (1.1). Multiplying the resulting equation by the complex conjugate $\Psi_{e,l}^{\star}$ of the electronic wave functions and integrating over the spatial coordinates $\mathbf{r}$ of the electrons gives the following set of differential equations:

$$\left(-\sum_{I=0}^{N-1}\frac{\hbar^2}{2M_I}\nabla_I^2 + E_l\right)\Psi_{n,l} + \sum_k \hat{C}_{lk}\Psi_{n,k} = i\hbar\frac{\partial}{\partial t}\Psi_{n,l}. \tag{2.18}$$

Here, $\hat{C}_{lk}$ is the nonadiabatic coupling operator which is defined as [6]:

$$\begin{aligned}\hat{C}_{lk} = &\int \Psi_{e,l}^{\star}\left(-\sum_{I=0}^{N-1}\frac{\hbar^2}{2M_I}\nabla_I^2\right)\Psi_{e,k}d\mathbf{r}\\ &+\sum_{I=0}^{N-1}\frac{1}{M_I}\left(\int \Psi_{e,l}^{\star}\left[-i\hbar\nabla_I\right]\Psi_{e,k}d\mathbf{r}\right)\left[-i\hbar\nabla_I\right].\end{aligned} \tag{2.19}$$

The goal of this operator is to account for the interactions between the multiple electronic and nuclear wave functions.

Afterwards, the set of differential equations given in eq. (2.18) can be decoupled. This is achieved by assuming that the effects of the coupling operator on the time evolution of Ψ are negligible, which gives:

$$\left(-\sum_{I=0}^{N-1}\frac{\hbar^2}{2M_I}\nabla_I^2 + E_l\right)\Psi_{n,l} = i\hbar\frac{\partial}{\partial t}\Psi_{n,l} \quad . \tag{2.20}$$

These decoupled differential equations correspond to the Born-Oppenheimer approximation [64]. This decoupling allows Ψ to be approximated as the product of an electronic and a nuclear wave function

$$\Psi\left(\mathbf{r},\mathbf{R},t\right) \approx \Psi_{e,l}\left(\mathbf{r},\mathbf{R}\right)\Psi_{n,l}\left(\mathbf{R},t\right) \quad . \tag{2.21}$$

The assumption that the coupling term is negligible is valid as long as the energy gap between different electronic wave functions remains small. This is not the case though in reactions that involve electron transfer for instance [65].

The last step in the derivation of the governing equations is to assume that the nuclei behave as classical point particles. According to the correspondence principle, the more massive a particle is, the better it will follow the laws of classical mechanics [66]. Since a proton is almost 2000 times the mass of an electron [63], this approximation is often valid. However, this approach can break down if effects such as zero-point vibrations and quantum tunneling of protons play an important role in the dynamics of a molecular system [6]. This is the case for hydrogen-bonded systems, such as those involving water, ice, or enzymes, where proton transfer affects properties like reaction rates [67].

2.4 Summary

Overall, applying the previous approximations leads to the following set of equations:

$$M_I \ddot{\mathbf{R}}_I\left(t\right) = -\nabla_I \langle \Psi_{e,0} | \hat{H}_e | \Psi_{e,0} \rangle \tag{2.22}$$

$$E_0 \Psi_{e,0} = \hat{H}_e \Psi_{e,0}, \tag{2.23}$$

where $\ddot{\mathbf{R}}_I$ is the acceleration of the Ith nucleus, $\Psi_{e,0}$ is the wave function of the electronic ground state and E_0 is the corresponding eigenenergy. These equations form the core of the Born-Oppenheimer molecular dynamics method. The purpose of the first one is to determine the classical nuclear motion. The right-hand side of that equation corresponds to the forces acting on the nuclei. These forces are found by computing the negative of the gradient of the potential energy surface (PES) $\langle \Psi_{e,0} | \hat{H}_e | \Psi_{e,0} \rangle$. They include all forces involving electrons as well as the repulsion forces between the nuclei. To determine the PES, the second equation is used. This equation is the time-independent electronic Schrödinger equation. The interested reader can refer to Refs. [6, 68] for a detailed derivation of the equations defining the BOMD method.

Since the goal of this research project is to harness the power of quantum computers to run molecular dynamics simulations, the PES was determined at every time step using the variational quantum eigensolver (VQE) [53]. This is an algorithm that takes advantage of both classical and quantum computers. Once the PES is known, its gradient is calculated on a classical computer to determine the forces acting on the nuclei. Afterwards, the position of all nuclei for the following time step is computed by numerically integrating the equations of motion associated with eq. (2.22). This process is repeated for all subsequent iterations [6]. A flowchart of the BOMD method, where the VQE is used to compute energy values on the PES, is shown in Fig. 2–2.

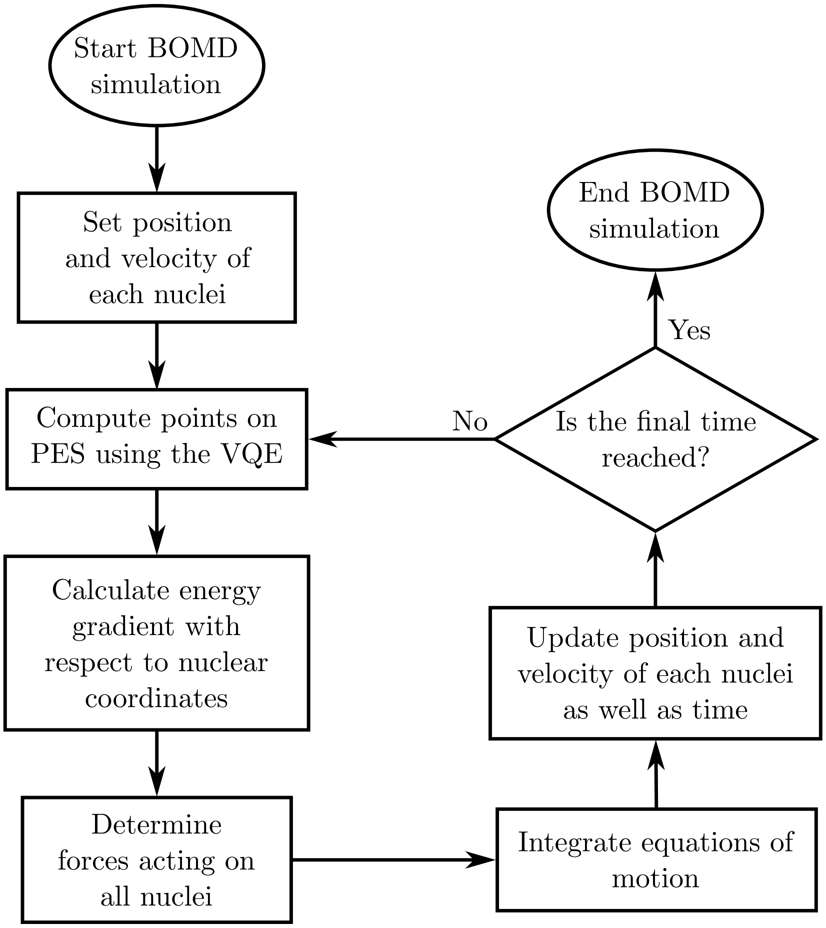

Figure 2–2: A flowchart of the BOMD method used for this research project.

CHAPTER 3
Variational Quantum Eigensolver

3.1 Overview

The variational quantum eigensolver (VQE) was presented by Peruzzo *et al.* [53] in 2014 as a method for solving the time-independent Schrödinger equation in polynomial time with the help of quantum computers. Even though the VQE takes advantage of the computational power of quantum computing, most of the steps in the algorithm take place on a classical computer. A flowchart showing the major tasks that must be undertaken is shown in Fig. 3–1.

The first step of the VQE is to select a reference quantum state $|\Psi_{ref}\rangle$ based on the position of each nuclei in a molecular system. This is achieved on a classical computer by using techniques such as the Hartree-Fock method [7] or the multiconfigurational self-consistent field (MCSCF) method [69, 70]. A unitary operator $U(\boldsymbol{\theta})$ known as an ansatz or an entangler is then selected. The goal of the ansatz is to generate trial states $|\Psi_t\rangle$ given some parameters $\boldsymbol{\theta}$. A trial state is created by applying the ansatz to the reference state, as shown in the following equation:

$$|\Psi_t\rangle = U(\boldsymbol{\theta})\,|\Psi_{ref}\rangle. \tag{3.1}$$

At this point, the ansatz and the electronic Hamiltonian are expressed in terms of fermionic creation and annihilation operators. These operators cannot be used directly by a quantum computer. Instead, they must be converted to Pauli operators which are suitable for

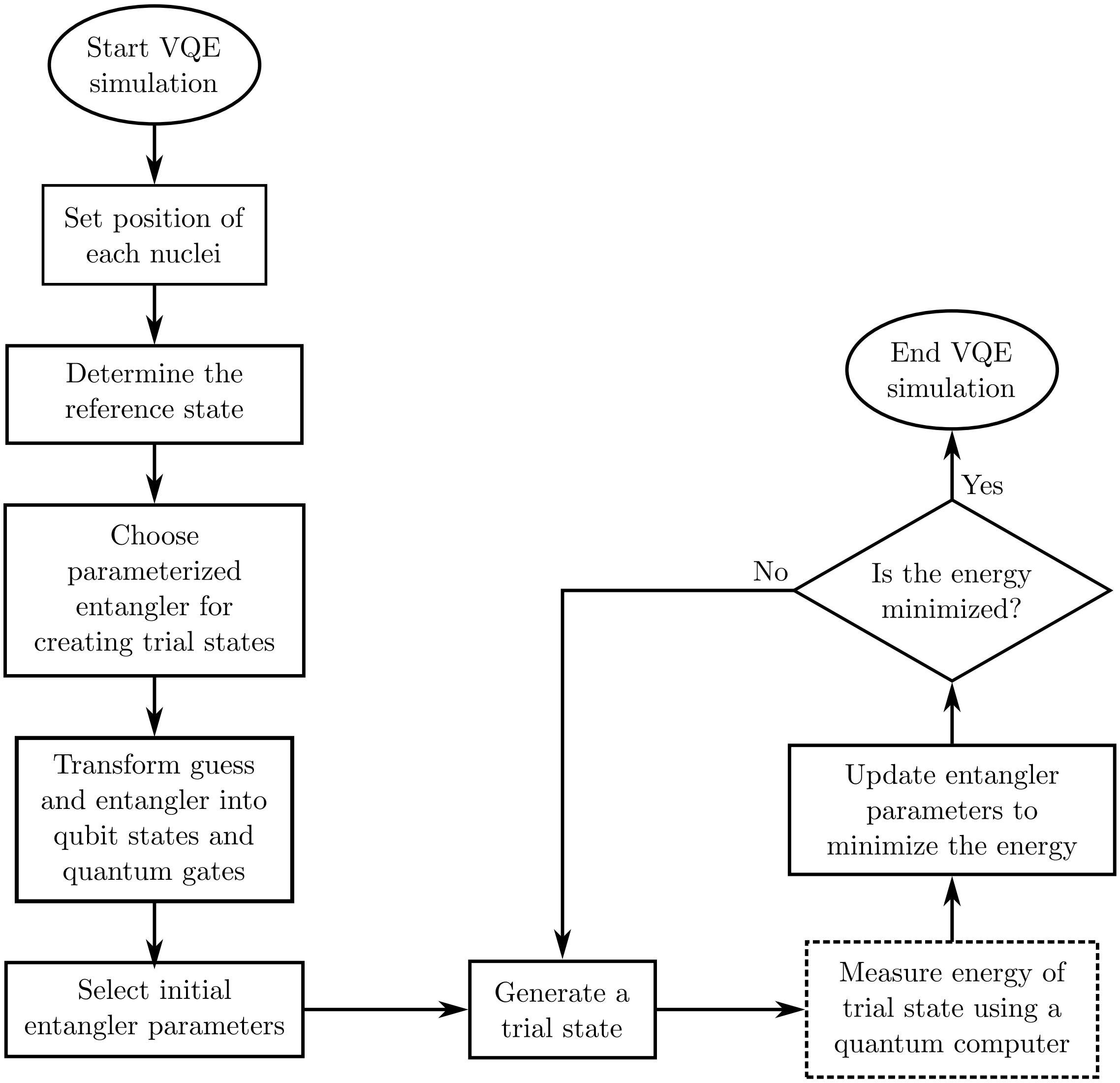

Figure 3–1: A flowchart of the VQE algorithm. The step in a dashed box is completed using a quantum computer while the others are completed using a classical computer.

quantum computation. $U(\boldsymbol{\theta})$ and $\hat{H}_e$ can then be transformed into quantum gates and qubit measurements [53]. More details regarding the fermionic operators will be given in Section 3.3, while methods for converting fermonic operators into quantum gates will be presented in Sections 3.4 and 3.5.

For the VQE's following step, the expectation value of the electronic Hamiltonian $\langle \hat{H}_e \rangle$ for a trial state is determined using a quantum computer. This expectation value corresponds to a point on the potential energy surface. To approximate it, a technique known as Hamiltonian averaging is employed [45]. First, the computer's qubits are initialized such that they are in the state $|\Psi_{ref}\rangle$. Then, a series of gates which implement $U(\boldsymbol{\theta})$ are applied to transform $|\Psi_{ref}\rangle$ into $|\Psi_t\rangle$. Afterwards, $\langle \hat{H}_e \rangle$ is estimated by measuring $|\Psi_t\rangle$ in an appropriate way. This procedure, from initializing the qubits as $|\Psi_{ref}\rangle$ to measuring $|\Psi_t\rangle$, is repeated multiple times to improve the accuracy of $\langle \hat{H}_e \rangle$. Mathematically, determining the total electronic energy E associated with the trial state $|\Psi_t\rangle$ can be expressed as:

$$E = \langle \hat{H}_e \rangle = \langle \Psi_t | \hat{H}_e | \Psi_t \rangle = \langle \Psi_{ref} | U^\dagger(\boldsymbol{\theta})\, \hat{H}_e U(\boldsymbol{\theta})\, | \Psi_{ref} \rangle, \tag{3.2}$$

where $\langle \Psi_{ref}|$ and $U^\dagger$ are the conjugate transpose of $|\Psi_{ref}\rangle$ and U, respectively. The Hamiltonian averaging procedure will be explained in more details in Section 3.6.

Once E is determined, it is sent along with the parameters $\boldsymbol{\theta}$ to an optimization algorithm on a classical computer. The optimizer then determines a new set of parameters in an attempt to minimize E. Those parameters are utilized to modify the ansatz and create a new trial state. Next, a new value for the energy is obtained on a quantum computer by using Hamiltonian averaging. This process is repeated until E is minimized [53]. It is important to note that the electronic TISE follows the variation principle [12]

$$\frac{\langle\Psi|\hat{H}_e|\Psi\rangle}{\langle\Psi|\Psi\rangle} \geq E_0, \tag{3.3}$$

where E_0 is the true ground state eigenenergy of $\hat{H}_e$. As a result, the difference between E and E_0 diminishes as lower values of E are determined. The optimization process that is employed to minimize E will be explained more thoroughly in Section 3.7.

3.2 Reference State

The first step in the VQE is to construct a reference wave function Ψ_{ref}. The wave function must describe all the electrons in the molecular system. A convenient way to do so is to use spin-orbitals ψ. Each spin-orbital can be occupied by an electron. One spin-orbital is the product of a spatial orbital $\phi(\mathbf{r})$, where $\mathbf{r}$ are spatial coordinates, and a spin eigenfunction $\sigma(m_s)$, where m_s is a discrete spin coordinate. The two possible spin eigenfunctions are defined as:

$$\alpha(m_s) = \begin{cases} 1 & \text{if } m_s = \frac{1}{2} \\ 0 & \text{if } m_s = -\frac{1}{2} \end{cases}, \tag{3.4}$$

and

$$\beta(m_s) = \begin{cases} 0 & \text{if } m_s = \frac{1}{2} \\ 1 & \text{if } m_s = -\frac{1}{2} \end{cases}. \tag{3.5}$$

Here, α is associated with a spin-up electron while β is associated with a spin-down one [12]. Since electrons in a molecule are indistinguishable fermions, they must satisfy the following antisymmetry requirement [4]:

$$\Psi_e\left(\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_{N_{SO}-1}\right) = -\Psi_e\left(\mathbf{x}_1, \mathbf{x}_0, \ldots, \mathbf{x}_{N_{SO}-1}\right), \tag{3.6}$$

where only the coordinates of electrons 0 and 1 have been interchanged and N_{SO} is the number of spin-orbitals used to construct the electronic wave function. Additionally, $\mathbf{x}_0$ includes both the spatial and spin coordinates of electron 0. For a system of two electrons, an electronic wave function containing two spin-orbitals can be written as:

$$\Psi_e\left(\mathbf{x}_0, \mathbf{x}_1\right) = \frac{1}{\sqrt{2}}\left(\psi_0\left(\mathbf{x}_0\right)\psi_1\left(\mathbf{x}_1\right) - \psi_0\left(\mathbf{x}_1\right)\psi_1\left(\mathbf{x}_0\right)\right). \tag{3.7}$$

As a consequence, it is not possible to have electrons that have the same spin and the same spatial orbital. This is known as the Pauli exclusion principle [71].

In order to enforce the antisymmetry requirement when constructing a wave function, a Slater determinant can be used [72]. For a system containing n electrons, a possible $n \times n$ Slater determinant ψ_{SD} is

$$\psi_{SD} = \frac{1}{\sqrt{n!}} \begin{vmatrix} \psi_0\left(\mathbf{x}_0\right) & \psi_1\left(\mathbf{x}_0\right) & \cdots & \psi_{N_{SO}-1}\left(\mathbf{x}_0\right) \\ \psi_0\left(\mathbf{x}_1\right) & \psi_1\left(\mathbf{x}_1\right) & \cdots & \psi_{N_{SO}-1}\left(\mathbf{x}_1\right) \\ \vdots & \vdots & \ddots & \vdots \\ \psi_0\left(\mathbf{x}_{n-1}\right) & \psi_1\left(\mathbf{x}_{n-1}\right) & \cdots & \psi_{N_{SO}-1}\left(\mathbf{x}_{n-1}\right) \end{vmatrix}. \tag{3.8}$$

It should be emphasized that a single Slater determinant may not be sufficient to describe the exact wave function Ψ_{exact} that satisfies the electronic TISE. To better approximate Ψ_{exact}, a linear combination of Slater determinants can be taken. This is shown in the following equation:

$$\Psi_{exact} \approx \sum_i a_i \left(\psi_{SD,i}\right), \tag{3.9}$$

where a_i are expansion coefficients. It is also important to note that the number of spin-orbitals must be greater than the number of electrons when multiple Slater determinants are needed [12]. In practice though, accurately approximating Ψ_{exact} is difficult because the number of possible Slater determinants grows factorially with the number of spin-orbitals [73]. Fortunately, the reference state used in the VQE does not need to be a highly accurate approximation of Ψ_{exact}.

For this research project, a method known as restricted Hartree-Fock (RHF) [7] was used to obtain the reference state. This approach works by finding the spin-orbitals that minimize $\langle \Psi_{RHF} | \hat{H}_e | \Psi_{RHF} \rangle$, where Ψ_{RHF} contains a single Slater determinant. The method is referred as restricted since the spatial orbitals are forced to be the same for both α and β spin-orbitals. This ensures that, similar to the exact wave function, Ψ_{RHF} is an eigenstate of the total and projected spin operator [12]. This is not the case for unrestricted Hartree-Fock where the spin-orbitals are allowed to differ spatially regardless of their spin [74].

Before proceeding any further, some terminology needs to be introduced. To start, the spatial orbitals which are optimized in the RHF method are known as molecular orbitals (MOs) and are denoted as ϕ_i. Each molecular orbital is composed of a linear combination of atomic orbitals (AOs), which are labeled as χ_j. This can be written as:

$$\phi_i = \sum_j C_{ji} \chi_j, \tag{3.10}$$

where C_{ji} are the coefficients that need to be optimized [7]. Those atomic orbitals are in turn made up of a linear combination of functions such as Gaussians orbitals. When these Gaussian orbitals are used, the AOs are known as contracted Gaussian-type orbitals (CGTOs). The generalized form of a CGTO which is centered at point A is shown in the following equation:

$$\chi = \sum_{\mu} d_{\mu} N_{\mu} \left(x_A^{n_G}\right) \left(y_A^{l_G}\right) \left(z_A^{m_G}\right) \exp\left(-\alpha_{\mu} r_A^2\right). \tag{3.11}$$

Here, N_μ is a normalization constant, α_μ and d_μ are scaling coefficients, r_A is the norm of the vector from point A to $\mathbf{r}$, while x_A, y_A, and z_A are the components of that vector. In addition, the shape of a CGTO is greatly influenced by the exponents n_G, l_G, and m_G. For instance, the sum of these exponents determines whether a CGTO is an s, p, or d orbital [75].

At this point, it is important to note that the atomic orbitals do not vary when employing the Hartree-Fock method. They are instead taken from what is called a basis set. Each basis set provides a set of equations with coefficients such as α_μ and d_μ. These coefficients are determined by optimizing the AOs so they minimize the energy of a single atom. An example is the STO-3G basis set that was created by Hehre *et al.* [76] in 1969, where each atomic orbital is represented by taking a linear combination of three Gaussian-type orbitals.

Once a basis set is selected, the molecular orbitals are optimized by solving the Roothaan equations [7]. These equations can be written as follows [7, 12]:

$$\mathbf{FC} = \mathbf{SC}\varepsilon, \tag{3.12}$$

where $\mathbf{F}$ is the Fock matrix, $\mathbf{C}$ is the matrix containing the MO coefficients, $\mathbf{S}$ is the overlap matrix, and ε is a diagonal matrix containing orbital energies. The Fock matrix depends not only on the kinetic and potential energy of atomic orbitals but also on the MO coefficients.

To be precise, each element of $\mathbf{F}$ is defined as:

$$F_{ij} = h_{ij}^{AO} + \sum_{k,l,m=0}^{N_{AO}-1} C_{km} C_{lm} \left(2g_{ijkl}^{AO} - g_{ilkj}^{AO}\right), \tag{3.13}$$

where h_{ij}^{AO} and g_{ijkl}^{AO} are one- and two-electron integrals, respectively, N_{AO} is the number of atomic orbitals, and C_{km} is an element of matrix $\mathbf{C}$. The integrals are determined using the following equation:

$$h_{ij}^{AO} = \int \chi_i^{\star}\left(\mathbf{r}\right) \left(-\frac{1}{2}\nabla^2 - \sum_{I=0}^{N-1} \frac{Z_I}{\left|\mathbf{R}_I - \mathbf{r}\right|}\right) \chi_j\left(\mathbf{r}\right) d\mathbf{r}, \tag{3.14}$$

and

$$g_{ijkl}^{AO} = \int\int \frac{\chi_i^{\star}\left(\mathbf{r}_1\right) \chi_k^{\star}\left(\mathbf{r}_2\right) \chi_j\left(\mathbf{r}_1\right) \chi_l\left(\mathbf{r}_2\right)}{\left|\mathbf{r}_1 - \mathbf{r}_2\right|} d\mathbf{r}_1 d\mathbf{r}_2, \tag{3.15}$$

where AO indicates that the integrals are computed in the basis of atomic orbitals and $\chi_i^{\star}$ is the complex conjugate of χ_i. It must also be pointed out that atomic units are used in eqs. (3.14) and (3.15). This means that the electron mass m_e, the elementary charge e, the reduced Planck constant $\hbar$, and the Coulomb constant $1/\left(4\pi\epsilon_0\right)$ were set to one [4]. As a result, the first term in the one-electron integral is the kinetic energy of an electron, while the second term is the Coulomb attraction energy between an electron and all the nuclei. Similarly, the two-electron integral is needed to compute the interelectronic Coulomb repulsion energy [12]. It is also important to note that the Fock matrix is not related to the total electronic energy itself, but rather to the variation in total electronic energy [77]. As for the overlap matrix $\mathbf{S}$, it appears in the Roothaan equations since the molecular orbitals are constrained to be orthonormal [78]. Each of its elements S_{ij} are computed using the integral

$$S_{ij} = \int \chi_i^{\star} \chi_j d\mathbf{r}. \tag{3.16}$$

This integral along with those given in eqs. (3.14) and (3.15) can be obtained using a method such as the McMurchie-Davidson scheme [75].

Once all the integrals are computed, it is then possible to solve the Roothaan equations. Although eq. (3.12) appears to be a generalized eigenvalue problem, that is not the case due to the presence of MO coefficients in the Fock matrix. An iterative procedure is therefore needed to determine **C** and ε. The procedure may go as follows [7]:

1. Make a guess for matrix **C**. This can be done by neglecting the term containing the two-electron integrals for the Fock matrix and then solving eq. (3.12).
2. Update the Fock matrix by substituting the new values for **C** into eq. (3.13).
3. Solve eq. (3.12) using the updated Fock matrix.
4. Repeat steps 2 and 3 until convergence is reached.

In this approach, values of **C** from only a single iteration are included when updating **F**. This can make convergence difficult or it may even lead to divergence. To prevent this issue, it is possible to take into account values of **C** from multiple iterations by employing a method such as the direct inversion in the iterative subspace (DIIS) [79, 80]. The result of both approaches is a matrix **C** where the columns correspond to the coefficients for a single molecular orbital. Each molecular orbital is associated with a spin-up molecular spin-orbital (MSO) and a spin-down MSO.

Since there are typically more MOs than there are electrons in the system, the MOs which contribute to the total energy must be determined. To do so, the MOs are divided into two categories based on the orbital energies. The $n/2$ molecular orbitals, ϕ_i, that have the lowest orbital energies ε_{ii}, are classified as occupied MOs. The rest of the MOs are defined as virtual. They are also sometimes known as unoccupied MOs [4].

With the MSOs divided into four categories based on spin and electron occupancy, it is now possible to define the reference state. For the VQE, $|\Psi_{ref}\rangle$ is defined according to the second-quantization language. In this language, each entry of a state vector represents an MSO. If an MSO is occupied, its entry is set to 1. Otherwise, it is set to 0. How the MSOs are ordered in a state vector is arbitrary however. For this research project, it was decided to place the entries in the following order:

$$|\Psi_{ref}\rangle = |\mathbf{k}_{\alpha,occ}\mathbf{k}_{\alpha,virt}\mathbf{k}_{\beta,occ}\mathbf{k}_{\beta,virt}\rangle. \tag{3.17}$$

Here, $\mathbf{k}$ represents a string of occupation numbers (0 or 1) associated with a set of MSOs. The occupation numbers $\mathbf{k}_{\alpha,occ}$ are for the occupied spin-up MSOs, while those for the virtual spin-down MSOs are represented by $\mathbf{k}_{\beta,virt}$. For example, the reference state for H_2 in a basis set containing 1 atomic orbital per hydrogen atom is:

$$|\Psi_{ref}\rangle = |1010\rangle, \tag{3.18}$$

where the first two entries are associated with spin-up MSOs. For the interested reader, detailed introductions to second quantization are given in Refs. [12] and [81].

It should be stressed that knowing $|\Psi_{ref}\rangle$ is not sufficient for fully characterizing the reference state when using the second-quantization language. It is also essential to define the electronic Hamiltonian. In this alternative formulation, $\hat{H}_e$ is represented using fermionic annihilation and creation operators, a_i and $a_i^\dagger$, respectively. The effect of these operators depends greatly on the occupancy number of entry i in a state vector. Once again, the interested reader should refer to Refs. [12] and [81] for an introduction to these fermionic operators. With them, $\hat{H}_e$ can be written as [12]:

$$\hat{H}_e = \sum_{P,Q=0}^{N_{MSO}-1} h_{PQ}^{MSO} a_P^\dagger a_Q + \frac{1}{2} \sum_{P,Q,R,S=0}^{N_{MSO}-1} g_{PQRS}^{MSO} a_P^\dagger a_R^\dagger a_S a_Q + h_{nuc}, \tag{3.19}$$

where N_{MSO} is the number of MSOs. The one-electron integral h_{PQ}^{MSO} is in the MSO basis and it corresponds to:

$$h_{PQ}^{MSO} = \int \sigma_P^\star\left(m_s\right) h_{pq}^{MO} \sigma_Q\left(m_s\right) dm_s \tag{3.20}$$

where $\sigma_P\left(m_s\right)$ is a spin eigenfunction. The first $N_{MSO}/2$ eigenfucntions are spin-up while the remainder are spin-down. As for the one-electron integral in the MO basis, it is defined as:

$$h_{pq}^{MO} = \int \phi_p^\star\left(\mathbf{r}\right) \left(-\frac{1}{2}\nabla^2 - \sum_{I=0}^{N-1} \frac{Z_I}{\left|\mathbf{R}_I - \mathbf{r}\right|}\right) \phi_q\left(\mathbf{r}\right) d\mathbf{r}, \tag{3.21}$$

where p is related to P according to the following relationship:

$$p = \begin{cases} P & \text{if } P < \frac{N_{MSO}}{2} \\ P - \frac{N_{MSO}}{2} & \text{otherwise} \end{cases}. \tag{3.22}$$

The index q is related to Q in a similar way. On the other hand, the two-electron integral in the MSO basis can be written as:

$$g_{PQRS}^{MSO} = \int\int \sigma_P^\star\left(m_{s,1}\right) \sigma_R^\star\left(m_{s,2}\right) g_{pqrs}^{MO} \sigma_Q\left(m_{s,1}\right) \sigma_S\left(m_{s,2}\right) dm_{s,1} dm_{s,2}, \tag{3.23}$$

where the two-electron integral in the MO basis corresponds to:

$$g_{pqrs}^{MO} = \int\int \frac{\phi_p^\star\left(\mathbf{r}_1\right) \phi_r^\star\left(\mathbf{r}_2\right) \phi_q\left(\mathbf{r}_1\right) \phi_s\left(\mathbf{r}_2\right)}{\left|\mathbf{r}_1 - \mathbf{r}_2\right|} d\mathbf{r}_1 d\mathbf{r}_2. \tag{3.24}$$

Once again, the indices r and s are defined in a similar way to p. As for h_{nuc}, it is the internuclear Coulomb repulsion energy. It can be calculated in atomic units using the following equation:

$$h_{nuc} = \frac{1}{2} \sum_{I,J=0,I\neq J}^{N-1} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|}, \tag{3.25}$$

where N is the number of nuclei in the system [12].

To obtain h_{pq}^{MO} and g_{pqrs}^{MO}, it is possible to utilize the integrals h_{ij}^{AO} and g_{ijkl}^{AO} that were computed in the RHF method. Assuming the MO coefficients are real, the matrix containing the integrals h_{pq}^{MO} corresponds to:

$$\mathbf{H}^{MO} = \mathbf{C}^T \mathbf{H}^{AO} \mathbf{C}, \tag{3.26}$$

where $\mathbf{C}^T$ is the transpose of $\mathbf{C}$. As for the integrals g_{pqrs}^{MO}, they can be computed by using the following equation:

$$g_{pqrs}^{MO} = \sum_{i,j,k,l=0}^{N_{AO}-1} C_{ip} C_{jq} C_{kr} C_{ls} g_{ijkl}^{AO}. \tag{3.27}$$

The previous equation can be applied efficiently by employing a scheme such as the one created by Nakata *et al.* [82].

Furthermore, eqs. (3.20) and (3.23), which are used to determine the integrals in the MSO basis, can be simplified by applying the spin orthonormality requirement [12]

$$\int \sigma_P^{\star}(m_s)\, \sigma_Q(m_s)\, dm_s = \delta_{\sigma_P \sigma_Q}, \tag{3.28}$$

where $\delta_{\sigma_P \sigma_Q}$ corresponds to the Kronecker delta. For the one-electron integrals, the result is:

$$h_{PQ}^{MSO} = \begin{cases} h_{pq}^{MO} & \text{if } \sigma_P = \sigma_Q \\ 0 & \text{otherwise} \end{cases}. \tag{3.29}$$

As a consequence, the only nonzero values of h_{PQ}^{MSO} are those involving two MSOs that have the same spin. As for the two-electron integrals, the simplification yields:

$$g_{PQRS}^{MSO} = \begin{cases} g_{pqrs}^{MO} & \text{if } \sigma_P = \sigma_Q \text{ and } \sigma_R = \sigma_S \\ 0 & \text{otherwise} \end{cases} . \tag{3.30}$$

It is important to note that σ_P does not necessarily have to be equal to σ_R for g_{PQRS}^{MSO} to be nonzero.

3.3 Parameterized Entangler

Once a reference state is determined, the next step in the VQE is to select a parameterized entangler $U(\boldsymbol{\theta})$. This entangler is sometimes referred to as an ansatz. It generates trial states $|\Psi_t\rangle$ that are made up of a linear combination of computational basis states [53]. On a two-qubit computer, the effect of applying the entangler to the reference state can be the following:

$$U(\boldsymbol{\theta})\,|\Psi_{ref}\rangle = a(\boldsymbol{\theta})\,|00\rangle + b(\boldsymbol{\theta})\,|01\rangle + c(\boldsymbol{\theta})\,|10\rangle + d(\boldsymbol{\theta})\,|11\rangle. \tag{3.31}$$

For the VQE to be tractable, the number of qubit operations required to apply the entangler must scale polynomially with the size of the system. In this case, the size of the system includes both the number of electrons and the number of MSOs. Furthermore, the entangler must be able to generate quantum states that contain a large number of computational basis states. To be more specific, the number of basis states must scale exponentially with the size of the system [48, 53]. This exponential growth is essential since the Hilbert space associated with the quantum state of a molecular system also scales exponentially.

In order to generate multiple basis states, the entangler must take advantage of entanglement. A set of qubits is said to be entangled when it cannot be expressed as a tensor product of single-qubit states [15]. For example, two qubits are entangled if they are in the state

$$|\Psi_{ent}\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right), \tag{3.32}$$

but they are not entangled if they are in the state

$$|\Psi_{sep}\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |10\rangle\right) = \frac{1}{\sqrt{2}} \left(|0\rangle + |1\rangle\right) \otimes |0\rangle. \tag{3.33}$$

One way to entangle qubits is to use the following entangling operator [15]:

$$U_{ent,c,t} = H_c U_{CN,c,t}, \tag{3.34}$$

where c is the control qubit, t is the target qubit, H_c is the Hadamard gate, and $U_{CN,c,t}$ is the two-qubit controlled-NOT gate. A mathematical description of these two gates is given in Section 3.5. The effect of applying the entangling operators to a two-qubit system, which is initially in the state $|00\rangle$, is the following:

$$U_{ent,0,1}|00\rangle = \frac{1}{\sqrt{2}} \left(|00\rangle + |11\rangle\right). \tag{3.35}$$

The result is a quantum state spanning two computational basis states. For a three-qubit system, two entangling operators can be applied to generate the quantum state

$$U_{ent,1,2}U_{ent,0,1}|000\rangle = \frac{1}{2} \left(|000\rangle + |011\rangle + |100\rangle - |111\rangle\right). \tag{3.36}$$

Entangled states can be created in a similar manner on quantum computers containing more qubits. For a system made up of m qubits, it is therefore possible to generate a quantum state spanning 2^{m-1} computational basis states using only $2 \times (m-1)$ quantum gates.

An entangler which can be used to produce trial states is the unitary coupled cluster (UCC) ansatz. This ansatz was proposed by Kutzelnigg [83] in 1977. It is defined as:

$$U_{UCC}\left(\boldsymbol{\theta}\right) = \exp\left(\hat{T}\left(\boldsymbol{\theta}\right) - \hat{T}^{\dagger}\left(\boldsymbol{\theta}\right)\right), \tag{3.37}$$

where $\hat{T}\left(\boldsymbol{\theta}\right)$ is a cluster operator. This operator can be written as:

$$\hat{T}\left(\boldsymbol{\theta}\right) = \sum_{i} \hat{T}_i\left(\boldsymbol{\theta}\right), \tag{3.38}$$

where $\hat{T}_i\left(\boldsymbol{\theta}\right)$ is an excitation operator. The purpose of each of these excitation operators is to include contributions from other configurations than the Hartree-Fock reference state. These configurations are generated by replacing a group of occupied orbitals in the reference sate by another group of virtual orbitals. For single and double excitations, these operators are defined as:

$$\hat{T}_1 = \sum_{\substack{i \in occ \\ k \in virt}} \theta_i^k a_k^{\dagger} a_i, \tag{3.39}$$

and

$$\hat{T}_2 = \sum_{\substack{i<j \in occ \\ k<l \in virt}} \theta_{ij}^{kl} a_k^{\dagger} a_l^{\dagger} a_i a_j. \tag{3.40}$$

Here, θ_{ij}^{kl} corresponds to an excitation amplitude where the virtual orbitals k and l have become occupied while the occupied orbitals i and j have become virtual [12]. In the same manner, it is also possible to define triple and quadruple excitation operators. However, the number of parameters in the VQE must scale polynomially with the size of the system. For this reason, it is common to truncate the UCC ansatz such that only single and double excitations are included. The result is the unitary coupled cluster singles-and-doubles (UCCSD) ansatz [84].

To be able to apply this entangler to the reference state using a quantum computer however, it is convenient to express it as a product of exponential functions. This can be accomplished by utilizing a technique known as trotterization [85]. When applied to the UCCSD ansatz, this technique generates the following approximation:

$$U_{UCCSD}\left(\boldsymbol{\theta}\right) = \exp\left[\sum_j \theta_j \left(\tau_j - \tau_j^\dagger\right)\right] \approx \left\{\prod_j \exp\left[\frac{\theta_j}{\rho}\left(\tau_j - \tau_j^\dagger\right)\right]\right\}^{\rho}, \tag{3.41}$$

where θ_j is an excitation amplitude, τ_j is a product of creation and annihilation operators, and ρ is known as the Trotter number [84]. Although the previous approximation becomes an equality when taking the limit as $\rho \to \infty$ [85], the Trotter number will be set to one for the remainder of this thesis. This will reduce the number of quantum gates needed to create the final state $|\Psi\rangle$. In addition, it is not necessary to apply the UCCSD ansatz exactly. The goal is only to modify $|\Psi_{ref}\rangle$ such that the energy associated with $\hat{H}_e$ is minimized.

After selecting an entangler, the next step is to pick initial values for the parameters $\boldsymbol{\theta}$. The option that will be chosen in this case is to use estimates for the excitation amplitudes from second-order Møller-Plesset perturbation theory (MP2) [9, 84]. These can be calulated using the following equations:

$$\theta_i^k = 0, \tag{3.42}$$

and

$$\theta_{ij}^{kl} = \frac{g_{likj}^{MSO} - g_{ljki}^{MSO}}{\epsilon_i + \epsilon_j - \epsilon_k - \epsilon_l}, \tag{3.43}$$

where ϵ_i is the orbital energy associated with the molecular spin-orbital i that was computed in the Hartree-Fock method.

3.4 Fermion-to-Qubit Mapping

So far, the electronic Hamiltonian and the parameterized ansatz have been defined in terms of fermionic annihilation and creation operators. However, these operators cannot be applied directly to the qubits on a quantum computer. Instead, they be must be converted into Pauli operators [15], which are defined as follows:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \text{and} \quad I = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.44}$$

To complete such a conversion, it is important to ensure that applying a qubit creation or annihilation operator has the same effect as applying a fermionic operator. For instance, the effect of "creating" an orbital i is the following [12]:

$$a_i^\dagger |k_0 \dots k_{i-1} 0 k_{i+1} \dots k_{N_{MO}-1}\rangle = (-1)^{\zeta_i} |k_0 \dots k_{i-1} 1 k_{i+1} \dots k_{N_{MO}-1}\rangle, \tag{3.45}$$

$$a_i^\dagger |k_0 \dots k_{i-1} 1 k_{i+1} \dots k_{N_{MO}-1}\rangle = 0, \tag{3.46}$$

where k_j is an occupation number and ζ_i is the parity, which is defined as:

$$\zeta_i = \sum_{j=0}^{i-1} k_j. \tag{3.47}$$

Similarly, applying a fermionic annihilation operator to a quantum state results in this outcome:

$$a_i |k_0 \dots k_{i-1} 1 k_{i+1} \dots k_{N_{MO}-1}\rangle = (-1)^{\zeta_i} |k_0 \dots k_{i-1} 0 k_{i+1} \dots k_{N_{MO}-1}\rangle, \tag{3.48}$$

$$a_i|k_0 \dots k_{i-1} 0 k_{i+1} \dots k_{N_{MO}-1}\rangle = 0. \tag{3.49}$$

Based on the previous equations, the properties that must be accounted for when applying a qubit creation or annihilation operator are the occupation number and the parity associated with an orbital. A method that was used during this research project to generate qubit operators that meet these requirements is the Jordan-Wigner transformation [86, 87]. Using this method, the state of a qubit determines if an orbital is occupied. For example, if a reference state is represented using the following occupation number vector:

$$|\Psi_{ref}\rangle = |1010\rangle, \tag{3.50}$$

then the quantum state which is stored by the qubits is also:

$$|\Psi_{ref}\rangle = |1010\rangle. \tag{3.51}$$

In matrix form, this state can be written as:

$$|\Psi_{ref}\rangle = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix} \otimes \begin{bmatrix} 0 \\ 1 \end{bmatrix} \otimes \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \tag{3.52}$$

where $\otimes$ stands for the Kronecker product.

Knowing that the qubits are in the occupation basis, all that must be done to change the occupation number of orbital i is to update qubit i. For a fermionic creation operator, the qubit creation operator $Q^\dagger$ must replace the qubit state by $|1\rangle$ if it encounters state $|0\rangle$, while it needs to set the state to 0 if it is applied to state $|1\rangle$. Mathematically, this is achieved by defining $Q^\dagger$ as [88, 89]:

$$Q^\dagger = |1\rangle\langle 0| = \frac{1}{2}(X - iY) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \tag{3.53}$$

In a similar fashion, the qubit annihilation operator Q is defined as:

$$Q = |0\rangle\langle 1| = \frac{1}{2}(X + iY) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}. \tag{3.54}$$

By themselves, Q and $Q^\dagger$ are not sufficient to represent the fermionic operators a and $a^\dagger$. Additional Pauli matrices must also be included to keep track of parity. They must be applied to all the qubits preceding qubit i and multiply the quantum state by -1 if an occupation number of 1 is encountered. The matrix that can achieve this task is Z. As a result, the fermionic creation and annihilation operators are equivalent to [88, 89]:

$$a_i^\dagger = Q_i^\dagger \bigotimes_{j<i} Z_j = \frac{1}{2}\left(X_i \bigotimes_{j<i} Z_j\right) - \frac{i}{2}\left(Y_i \bigotimes_{j<i} Z_j\right), \tag{3.55}$$

and

$$a_i = Q_i \bigotimes_{j<i} Z_j = \frac{1}{2}\left(X_i \bigotimes_{j<i} Z_j\right) + \frac{i}{2}\left(Y_i \bigotimes_{j<i} Z_j\right), \tag{3.56}$$

where $\bigotimes_{j<i} Z_j$ means that the matrix Z is applied to all qubits with index $j < i$. For example, on a quantum computer with 4 qubits, creation operator $a_2^\dagger$ is defined as:

$$a_2^\dagger = \frac{1}{2}(Z_0 \otimes Z_1 \otimes X_2 \otimes I_3) - \frac{i}{2}(Z_0 \otimes Z_1 \otimes Y_2 \otimes I_3) = \frac{1}{2}(Z_0 Z_1 X_2) - \frac{i}{2}(Z_0 Z_1 Y_2), \tag{3.57}$$

where I is the identity matrix.

Once the method for mapping fermionic operators into qubit operators is chosen, it becomes possible to express the electronic Hamiltonian in terms of Pauli operators. The result is a linear combination of tensor products of Pauli operators P_i. The electronic Hamiltonian can therefore be written as:

$$\hat{H}_e = \sum_i d_i P_i, \tag{3.58}$$

where d_i is a real coefficient. It is important to note that each tensor product can be simplified such that at most one Pauli operator is associated with each qubit. This is because the Pauli matrices along with I and the multiplicative factors ± 1 and $\pm i$ form a group. As a consequence, the product of two members of the Pauli group is another member of that group [15]. For example, the tensor product,

$$P = Z_1 Z_1 \otimes X_2 Y_2 \otimes Y_3, \tag{3.59}$$

can be simplified to:

$$P = iZ_2 \otimes Y_3. \tag{3.60}$$

After writing the electronic Hamiltonian in terms of Pauli operators, there still remains the task of transforming the parameterized entangler. In the case of the UCCSD ansatz, the result is a product of exponential operators containing entangler parameters θ_j and Pauli operators. The transformed ansatz is shown in the following equation:

$$U_{UCCSD}\left(\boldsymbol{\theta}\right) = \prod_j \exp\left(i\theta_j \sum_l c_{j,l} P_{j,l}\right), \tag{3.61}$$

where $c_{j,l}$ are real coefficients that arise when applying the transformation and $P_{j,l}$ are tensor products of Pauli operators. Additionally, the Trotter number was set to 1 in this case. However, the ansatz is not in its simplest form at this point. It is in fact possible to write the entangler as:

$$U_{UCCSD}\left(\boldsymbol{\theta}\right) = \prod_{j,l} \exp\left(i c_{j,l} \theta_j P_{j,l}\right). \tag{3.62}$$

This simplification is valid since, as Romero *et al.* [84] demonstrated, the tensor products of Pauli operators associated with the same parameters θ_j commute with each other. Mathematically, this means that:

$$\left[P_{j,l_1}, P_{j,l_2}\right] = P_{j,l_1} P_{j,l_2} - P_{j,l_2} P_{j,l_1} = 0. \tag{3.63}$$

Knowing that these operators commute, it is then possible to use the Baker-Campbell-Hausdorff formula to rewrite eq. (3.61) as eq. (3.62). This formula states that:

$$\exp\left(\mathbf{A}\right)\exp\left(\mathbf{B}\right) = \exp\left(\mathbf{A} + \mathbf{B} + \frac{1}{2}\left[\mathbf{A}, \mathbf{B}\right] + \dots\right), \tag{3.64}$$

where **A** and **B** are sufficiently small matrices and terms corresponding to "…" contain additional commutators involving **A** and **B** [90].

3.5 Quantum Circuit Construction

The implementation of an algorithm on a quantum computer is typically represented by a quantum circuit. Such a circuit shows in which order quantum gates are applied to the qubits on a quantum computer. An example of a quantum circuit is shown in Fig. 3–2. All the circuit diagrams shown in this thesis, including this one, were generated using IBM's Qiskit library [91].

The circuit of Fig. 3–2 contains two qubits, represented by the first two horizontal lines, as well as two classical bits, represented by the last two horizontal lines. The horizontal axis represents time. This means that the quantum gate which is the furthest to the left is applied first. In this case, the qubits are initially in the ground state $|00\rangle$. The first gate

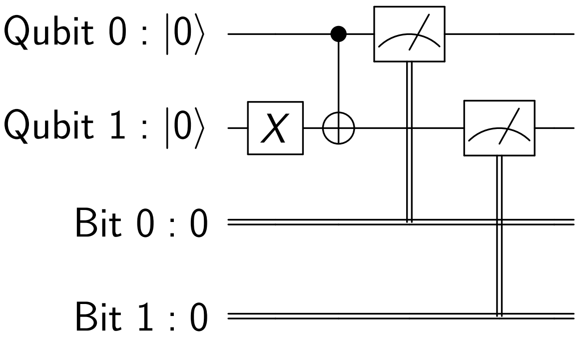

Figure 3–2: A sample quantum circuit made up of two qubits and two classical bits. It contains a Pauli-X gate, a CNOT gate, and two qubit measurements.

of the circuit is a Pauli-X gate that is applied to the second qubit. It is sometimes called the bit flip gate since it replaces $|0\rangle$ by $|1\rangle$ and vice versa [15]. After this gate, the quantum computer is in state $|01\rangle$.

The next gate that is applied in Fig. 3–2 is a two-qubit controlled-NOT or CNOT gate. The black dot in the gate's diagram is placed on the control qubit's line while the large circle is positioned on the target qubit's line. The effect of a CNOT gate is to apply a Pauli-X gate to the target qubit if the control qubit is $|1\rangle$, but to leave it unchanged if the control qubit is $|0\rangle$. Mathematically, it is described as:

$$|0\rangle\langle 0|_{control} \otimes I_{target} + |1\rangle\langle 1|_{control} \otimes X_{target}. \tag{3.65}$$

In matrix form, the CNOT gate shown in the previous figure corresponds to:

$$U_{CN,0,1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \tag{3.66}$$

where the first index in $U_{CN,0,1}$ is associated with the control qubit while the second one represents the target qubit [15].

The last elements of Fig. 3–2 are qubit measurements. These operations change the value of a classical bit based on the state of a measured qubit. The result of these measurements is probabilistic in nature. For example, if measurements are taken on a two-qubit quantum computer when it is in the following state:

$$|\Psi\rangle = \alpha|01\rangle + \beta|11\rangle, \tag{3.67}$$

the first classical bit will be set to 0 with probability $|\alpha|^2$ and to 1 with probability $|\beta|^2$. On the other hand, the second classical bit will always be set to 1. To determine the probability $p_{|\Omega\rangle}$ of measuring a state Ω, the next equation can be employed:

$$p_{|\Omega\rangle} = |\langle\Omega|\Psi\rangle|^2 = \langle\Omega|\Psi\rangle\langle\Psi|\Omega\rangle, \tag{3.68}$$

where $|\Psi\rangle$ is the state which was generated on the quantum computer [62]. For the sample circuit, the quantum computer's final state before the measurements is:

$$|\Psi\rangle = U_{CN,0,1}X_1|00\rangle = U_{CN,0,1}|01\rangle = |01\rangle. \tag{3.69}$$

This means that the first classical bit will always be set to 0 while the second one will always be set to 1.

In the case of the VQE, a quantum circuit is needed to generate the final state $|\Psi\rangle$. The first portion of this circuit is dedicated to creating the reference state $|\Psi_{ref}\rangle$. If this state is not in a superposition, it can be initialized by applying Pauli-X gates to qubits which need

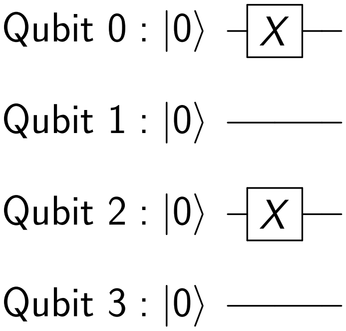

Figure 3–3: The quantum circuit used to initialize state $|1010\rangle$.

to be set to 1. For instance, if the reference state is $|1010\rangle$, it can be generated as follows:

$$|\Psi_{ref}\rangle = X_0 X_2 |0000\rangle = |1010\rangle, \tag{3.70}$$

where this circuit is presented graphically in Fig. 3–3.

After initializing the reference state, the next task is to apply the parameterized entangler. For an entangler that is based on the UCCSD ansatz, the quantum operation that needs to be applied was presented previously in eq. (3.62) and corresponds to:

$$U_{UCCSD}(\boldsymbol{\theta}) = \prod_{j,l} \exp\left(i c_{j,l} \theta_j P_{j,l}\right). \tag{3.71}$$

To understand how to generate the circuit for this operation, it is helpful to look at a single exponential term. For instance, a term in eq. (3.71) could be the following:

$$U_{j,l}(\theta_j) = \exp\left(i c_{j,l} \theta_j X_0 X_1 Y_2 X_3\right). \tag{3.72}$$

That term can be reexpressed as the product of a unitary operator U_{CB}, and an exponential function containing only Pauli-Z matrices,

$$U_{j,l}(\theta_j) = U_{CB}^{\dagger} \exp\left(i c_{j,l} \theta_j Z_0 Z_1 Z_2 Z_3\right) U_{CB}. \tag{3.73}$$

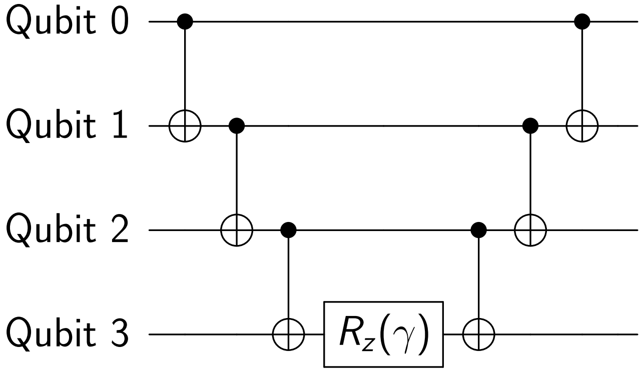

Figure 3–4: The quantum circuit that applies the operation $\exp\left(-i\frac{\gamma}{2}Z_0Z_1Z_2Z_3\right)$.

The exponential operator which is present in eq. (3.73) can be applied on a quantum computer by placing a z-rotation gate, $R_z(\gamma)$, in between series of CNOT gates, as shown in Fig. 3–4 [92]. The effect of the circuit given in Fig. 3–4 is to apply the following unitary operator:

$$U_{exp} = \exp\left(-i\frac{\gamma}{2}Z_0Z_1Z_2Z_3\right), \tag{3.74}$$

where γ corresponds to the rotation angle associated with the $R_z(\gamma)$ gate. This z-rotation gate can be written as [15]:

$$R_z(\gamma) = \exp\left(-i\frac{\gamma}{2}Z\right) = \begin{bmatrix} e^{-i\frac{\gamma}{2}} & 0 \\ 0 & e^{i\frac{\gamma}{2}} \end{bmatrix}. \tag{3.75}$$

For the exponential operator of eq. (3.73), the rotation angle must be set to

$$\gamma = -2c_{j,l}\theta_j. \tag{3.76}$$

At this stage, all that remains for creating the quantum circuit associated with eq. (3.72) is to determine the quantum gates for the operators U_{CB}. Their effect is to change the basis

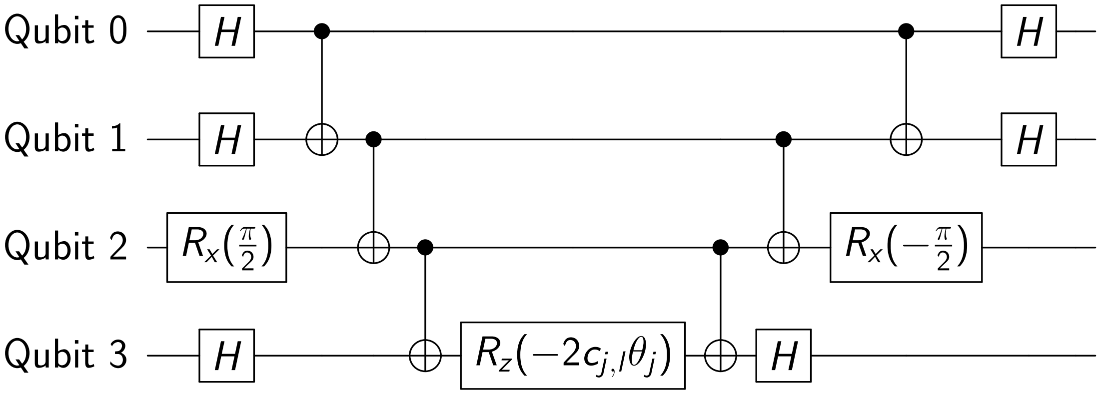

Figure 3–5: The quantum circuit that applies the operation $\exp\left(ic_{j,l}\theta_j X_0 X_1 Y_2 X_3\right)$.

from Z to either X or Y. To change the basis from Z to X, a Hadamard gate H is employed. On the other hand, placing a qubit in the Y basis can be accomplished by using an x-rotation gate $R_x(\gamma)$ with an angle of $\frac{\pi}{2}$. The H and $R_x(\gamma)$ gates are defined as follow:

$$H = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \tag{3.77}$$

$$R_x(\gamma) = \exp\left(-i\frac{\gamma}{2}X\right) = \begin{bmatrix} \cos\frac{\gamma}{2} & -i\sin\frac{\gamma}{2} \\ -i\sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{bmatrix}. \tag{3.78}$$

After the CNOT and z-rotation gates are applied, it is also important to return the qubit to its original basis. To return from the X basis to the Z basis, another Hadamard gate can be used. For the change of basis from Y to Z however, an $R_x(\gamma)$ gate with an angle of $-\frac{\pi}{2}$ must be utilized [92]. As a result, the exponential operator of eq. (3.72) can be applied using the quantum circuit shown in Fig. 3–5. This means that the circuit for the ansatz of eq (3.71) corresponds to a series of sub-circuits, such as the one in Fig. 3–5, where each of these sub-circuits is associated with an exponential term.

3.6 Hamiltonian Averaging

In order to efficiently determine the energy which results from applying the electronic Hamiltonian $\hat{H}_e$ to the trial state $|\Psi_t\rangle$, a method known as Hamiltonian averaging is employed [53]. This method takes advantage of the fact that the Hamiltonian is a linear combination of Pauli operator products. This means that the expectation value $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle$ is equal to the linear combination of the expectation value of Pauli operator products $\langle\Psi_t|\hat{P}_i|\Psi_t\rangle$, as shown in the following equation:

$$\langle\Psi_t|\hat{H}_e|\Psi_t\rangle = \sum_i h_i \langle\Psi_t|\hat{P}_i|\Psi_t\rangle, \tag{3.79}$$

where the coefficients h_i are real.

By convention, reading out a qubit is completed in the basis of the Z operator. As a result, additional quantum gates must be applied to qubits where readout needs to be achieved in the X or Y basis. To perform a change of basis from Z to Y, an x-rotation gate with a $\frac{\pi}{2}$ angle is applied. Similarly, in the case where the qubit must be placed in the X basis, a y-rotation gate with an angle of $-\frac{\pi}{2}$ is used. This y-rotation gate $R_y(\gamma)$ is defined as [15]:

$$R_y(\gamma) = \exp\left(-i\frac{\gamma}{2}Y\right) = \begin{bmatrix} \cos\frac{\gamma}{2} & -\sin\frac{\gamma}{2} \\ \sin\frac{\gamma}{2} & \cos\frac{\gamma}{2} \end{bmatrix}. \tag{3.80}$$

Consequently, measuring the expectation value $\langle\Psi_t|X_0Z_1Y_2Z_3|\Psi_t\rangle$ can be achieved by using the circuit shown in Fig. 3–6. The result of each qubit readout is stored as a value of 0 or 1 in a classical bit. It is important to note however that a single measurement is not sufficient to determine an expectation value. To compute the expectation value

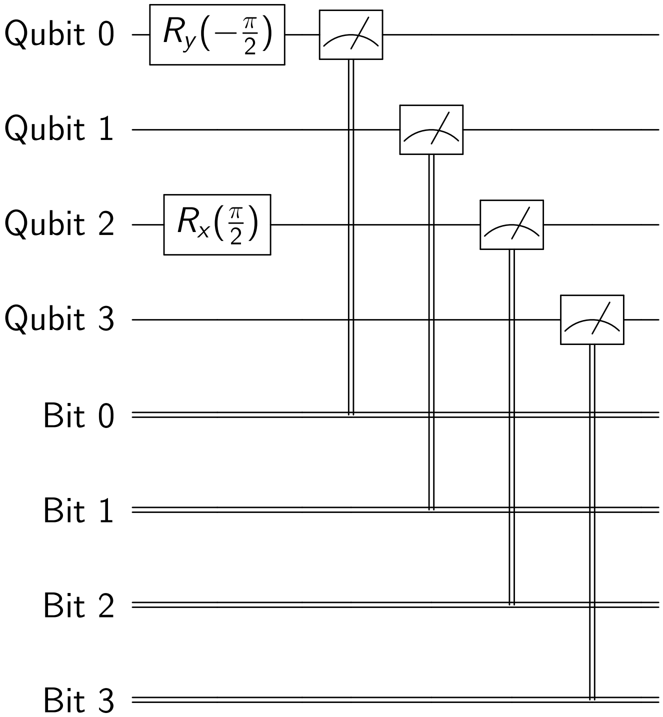

Figure 3–6: The quantum circuit used to measure the expectation value $\langle \Psi_t | X_0 Z_1 Y_2 Z_3 | \Psi_t \rangle$.

$\langle \Psi_t | Z_0 Z_1 \dots Z_{N_r-2} Z_{N_r-1} | \Psi_t \rangle$, the following equation can be used:

$$\langle \Psi_t | Z_0 Z_1 \dots Z_{N_r-2} Z_{N_r-1} | \Psi_t \rangle = \sum (-1)^{\eta} \, p_{|b_0 b_1 \dots b_{N_r-2} b_{N_r-1}\rangle}, \tag{3.81}$$

where $p_{|b_0 b_1 \dots b_{N_r-2} b_{N_r-1}\rangle}$ is the probability of measuring state $|b_0 b_1 \dots b_{N_r-2} b_{N_r-1}\rangle$. The sum in this case is taken over every possible state. Additionally, N_r represents the number of qubit readouts and b_i can take a value of either 0 or 1. The exponent η is determined according to this equation:

$$\eta = \begin{cases} 0 & \text{if } \sum_{i=0}^{N_m-1} b_i \text{ is even} \\ 1 & \text{if } \sum_{i=0}^{N_m-1} b_i \text{ is odd} \end{cases}. \tag{3.82}$$

As a consequence of eq. (3.81), determining an expectation value requires taking multiple measurements in order to approximate the probability of measuring each state $|b_0 b_1 \dots b_{N_r-2} b_{N_r-1}\rangle$. As an example, if two qubits are measured on 100 occasions and the outcomes 00, 01, and 11 are obtained 29, 48, and 23 times, respectively, then the expectation value $\langle \Psi_t | Z_0 Z_1 | \Psi_t \rangle$ is approximated as follows:

$$\langle \Psi_t | Z_0 Z_1 | \Psi_t \rangle \approx p_{|00\rangle} - p_{|01\rangle} - p_{|10\rangle} + p_{|11\rangle} = 0.04. \tag{3.83}$$

It must be emphasized however that it is necessary to recreate the quantum state $|\Psi_t\rangle$ every time a new measurement needs to be taken. This is because the state collapses after it is measured. For instance, if the outcome of reading out a qubit which is initially in state $\alpha|0\rangle + \beta|1\rangle$ is 1, then taking any new readouts of that qubit will always yield a value of 1 [62]. This means that the steps of initializing $|\Psi_{ref}\rangle$ and applying $U_{UCCSD}(\mathbf{t})$ need to be repeated multiple times for approximating each $\langle \Psi_t | \hat{P}_i | \Psi_t \rangle$ in eq. (3.79).

Another important aspect of Hamiltonian averaging is to predict how many measurements are needed to calculate an expectation value with a certain uncertainty ε. If the total number of measurements required m_{req} is divided equally among the M nonconstant terms in eq. (3.79), the following number of measurements required can be used [54, 93]:

$$m_{req} = \left(\frac{Mz^2}{\varepsilon^2}\right)\sum_{i=1}^{M} h_i^2, \tag{3.84}$$

where z corresponds to the z score associated with the upper bound of the interval $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle \pm \varepsilon$. More details regarding this equation and how it was derived are given in the Appendix.

3.7 Energy Optimization

The last part of the VQE algorithm is the classical optimization scheme. Its purpose is to update the parameters $\boldsymbol{\theta}$ of the entangler such that the expectation value of the electronic Hamiltonian is minimized [53]. For this research project, the simultaneous perturbation stochastic approximation (SPSA) algorithm was used [94, 95]. This algorithm is well suited for the VQE since it requires a small number of objective function evaluations and it performs well in the presence of noise.

In order to determine new parameters $\boldsymbol{\theta}$, the SPSA algorithm uses the following approximation for the directional derivative:

$$\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right) = \frac{E\left(\boldsymbol{\theta}_k + \delta_k \hat{\mathbf{v}}_k\right) - E\left(\boldsymbol{\theta}_k - \delta_k \hat{\mathbf{v}}_k\right)}{2\delta_k}, \tag{3.85}$$

where E corresponds to the total electronic energy, $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle$, and δ_k is the derivative step size for iteration k. As for the direction vector $\mathbf{v}_k$, each of its entries is randomly chosen from a Bernoulli distribution where the outcomes 1 and -1 can be selected with an equal probability. It is important to note however that the normalized form $\hat{\mathbf{v}}_k$ of the direction vector is utilized. In the case of the step size, it is computed using the following equation:

$$\delta_k = c_k |\mathbf{v}_k|, \tag{3.86}$$

where $|\mathbf{v}_k|$ is the norm of the direction vector and c_k is the derivative gain sequence. The latter variable is defined as:

$$c_k = \frac{c}{(k+1)^{\gamma}}, \tag{3.87}$$

where k is set to 0 for the first iteration. The value for γ on the other hand is set to 0.101. This value is considered to be practically optimal by Spall [95]. As for c, it can be chosen by first noting that it is related to the step size according to:

$$c = \frac{\delta_0}{|\mathbf{v}_0|}. \tag{3.88}$$

The parameter c must therefore be selected based on the desired directional derivative step size for the first iteration. If the step size is increased, the error resulting from employing the finite difference method to approximate the directional derivative will be larger. On the other hand, if the step size is decreased, the number of measurements required to compute the directional derivative with a certain uncertainty rises.

Once the directional derivative is approximated, the parameters are updated by using this equation:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - a_k |\mathbf{v}_k| \left(\widetilde{\nabla}_{\mathbf{v}} E \left(\boldsymbol{\theta}_k \right) \right) \mathbf{v}_k^{-1}, \tag{3.89}$$

where a_k corresponds to the parameter updater gain sequence and $\mathbf{v}_k^{-1}$ is obtained by replacing each entry of the direction vector by its reciprocal. The gain sequence in this case is defined as:

$$a_k = \frac{a}{(A+k+1)^{\alpha}}, \tag{3.90}$$

where the optimal value of α is 0.602, while A and a are user-defined parameters. According to Spall [95], A should be set to approximately 10% of the number of expected iterations. As for a, it should be chosen based on the desired change in $\boldsymbol{\theta}$ for the initial step. Rewriting eq. (3.89) as:

$$\boldsymbol{\theta}_{k+1} = \boldsymbol{\theta}_k - s_k \hat{\mathbf{v}}_k^{-1}, \tag{3.91}$$

where s_k is related to the change in $\boldsymbol{\theta}$, allows a to be selected using the following equation:

$$a = \frac{s_0 \left(A+1\right)^{\alpha}}{|\mathbf{v}_0||\mathbf{v}_0^{-1}| \left(\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_0\right)\right)}. \tag{3.92}$$

The value of a will affect the rate of convergence of the SPSA algorithm. For example, it might require a large number of iterations for $\boldsymbol{\theta}$ to change appreciably if a is too small. On the other hand, choosing a value of a that is too large can cause the algorithm to repeatedly overestimate the value of s_k required to reach the optimal point.

After new parameters $\boldsymbol{\theta}_{k+1}$ are determined, k is updated and this process is repeated until some termination conditions are satisfied. A convenient choice in this scenario is to stop the optimization algorithm once the absolute value of the directional derivative is below a certain threshold. It is also important to emphasize that the success of this algorithm depends greatly on the accuracy of the values that are computed for the directional derivative. For this reason, an equation which determines the total number of energy measurements required $m_{d,req}$ to calculate $\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)$ with uncertainty ε_d was derived. The derivation was based on the approaches used by Wecker *et al.* [93] and McClean *et al.* [54] to determine eq. (3.84).

The equation for $m_{d,req}$ is the following:

$$m_{d,req} = \left(\frac{\left(z_d\right)^2 M}{\delta^2 \varepsilon_d^2} \right) \sum_{i=1}^{M} h_i^2, \tag{3.93}$$

where δ is the derivative's step size and z_d is the z score of the upper bound of the interval $\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right) \pm \varepsilon_d$. Additional details concerning this equation and how it was derived are given in the Appendix. Furthermore, this result is useful in choosing a value for c. The effect of c is to vary the step size in the directional derivative. If c is too small, the number of measurements required to reach a certain uncertainty will be very large. On the other hand, if c is too large, the accuracy of eq. (3.85) will decrease.

CHAPTER 4
Potential Energy Surface of Molecular Hydrogen

4.1 Implementation

Before implementing the entirety of the BOMD method, it was essential to first test the VQE. To do so, the expectation value of the electronic Hamiltonian was computed for molecular hydrogen (H_2) at different interatomic distances. To calculate these values, C++ libraries for almost every step of the VQE were created by the author of this thesis. For generating a reference state $|\Psi_{ref}\rangle$, the STO-3G basis set was employed [76]. This is considered a minimal basis set since it contains the minimum number of atomic orbitals required to accommodate the electrons of an atom. The advantage is that a small number of qubits are needed since the number of molecular orbitals generated is kept to a minimum. Another benefit is that the number of quantum gates required to entangle qubits in order to generate a trial state $|\Psi_t\rangle$ is also small. To integrate the atomic orbitals which are needed for the Hartree-Fock method, the McMurchie-Davidson scheme [75] was implemented. One of the only external libraries that was used to help create reference states was the Elemental library [96]. It was utilized to perform matrix operations and to solve the Roothaan equations, given in eq. (3.12).

Additional C++ libraries were created by the author to perform tasks such as generating parameterized entanglers, transforming fermionic operators into Pauli operators, creating quantum circuits, and optimizing the trial wave function. On the other hand, it was decided to use IBM's Qiskit library [91] in order to determine the expectation value of the electronic

Hamiltonian. This library contains two quantum computer simulators known as the state vector simulator and the QASM (Quantum Assembly Language) simulator. The former simulator works by applying quantum gates to an initial quantum state vector $|\Psi_i\rangle$ in order to generate a final quantum state vector $|\Psi_f\rangle$. The expectation value of a Hamiltonian $\hat{H}$ can then be computed exactly by performing the following matrix multiplication:

$$\langle \hat{H} \rangle = \langle \Psi_f | \hat{H} | \Psi_f \rangle. \tag{4.1}$$

On the other hand, the QASM simulator replicates the behavior of an ideal quantum computer. To obtain an expectation value, each quantum circuit must be applied and measured numerous times as per the Hamiltonian averaging procedure. Another feature of the Qiskit library is the ability to remotely access one of IBM's quantum computers. As of July 2019, they allow users to run quantum simulations on three of their computers. The largest one contains 14 qubits, while the other two have 5 qubits.

4.2 Potential Energy Surface

The first test was to calculate the total electronic energy $\langle \hat{H}_e \rangle$ at the equilibrium interatomic distance d_{eq} for molecular hydrogen. For this molecule, the experimental value of d_{eq} corresponds to 0.7414 Å [97]. For this test, Qiskit's state vector simulator was used. Coefficients a and c for the SPSA optimization algorithm were set to 0.7 and 0.001 respectively. For comparison, the energy $\langle \hat{H}_e \rangle$ was determined using the restricted Hartree-Fock (RHF) method with the author's code as well as with the full configuration interaction (FCI) method in Gaussian 16 [98]. The FCI method requires using all possible Slater determinants, given a certain basis set, when generating the electronic wave function [12]. For the three previous simulations, the STO-3G basis set was used. It should be noted that it is not

Table 4–1: The total electronic energy of H_2 at the equilibirum distance. The energy calculated using FCI with the aug-cc-pV6Z basis set was set as the reference for determining the energy error.

Method	RHF	VQE	FCI	FCI
Basis set	STO-3G	STO-3G	STO-3G	aug-cc-pV6Z
Energy ($\mathrm{E_h}$)	-1.1166	-1.1373	-1.1373	-1.1744
Energy error ($\mathrm{mE_h}$)	57.7	37.1	37.1	0.0

possible to determine the total electronic energy experimentally. As an alternative reference point, the FCI calculation was repeated in Gaussian 16 using a large basis set known as aug-cc-pV6Z [99, 100]. The results of these simulations are shown in Table 4–1. It is also important to note that the energy error was calculated with reference to the FCI calculation with the aug-cc-pV6Z basis set. Furthermore, energy values in Table 4–1 are reported in Hartree ($\mathrm{E_h}$).

As desired, the energy values are the same for both the VQE and FCI methods using the STO-3G basis set. This signifies that the UCCSD-based ansatz contained enough parameters to fully optimize the trial state $|\Psi_t\rangle$. However, the STO-3G basis set is not large enough to allow the results to be within chemical accuracy, which is defined as 1 kcal/mole (1.6 $\mathrm{mE_h}$). This target accuracy was set such that numerical results for atomization energies or heats of formations, for instance, would also be within experimental accuracy [101]. Additionally, it should be emphasized that all the methods that were used to compute the energy satisfy the variation principle of eq. (3.3). As a result, the values of energy error given in Table 4–1 are greater or equal to the exact energy error, which is calculated with respect to the exact energy.

To further test the validity of the VQE implementation, the potential energy curve was also obtained for H_2. The same methods as those used for calculating the total electronic

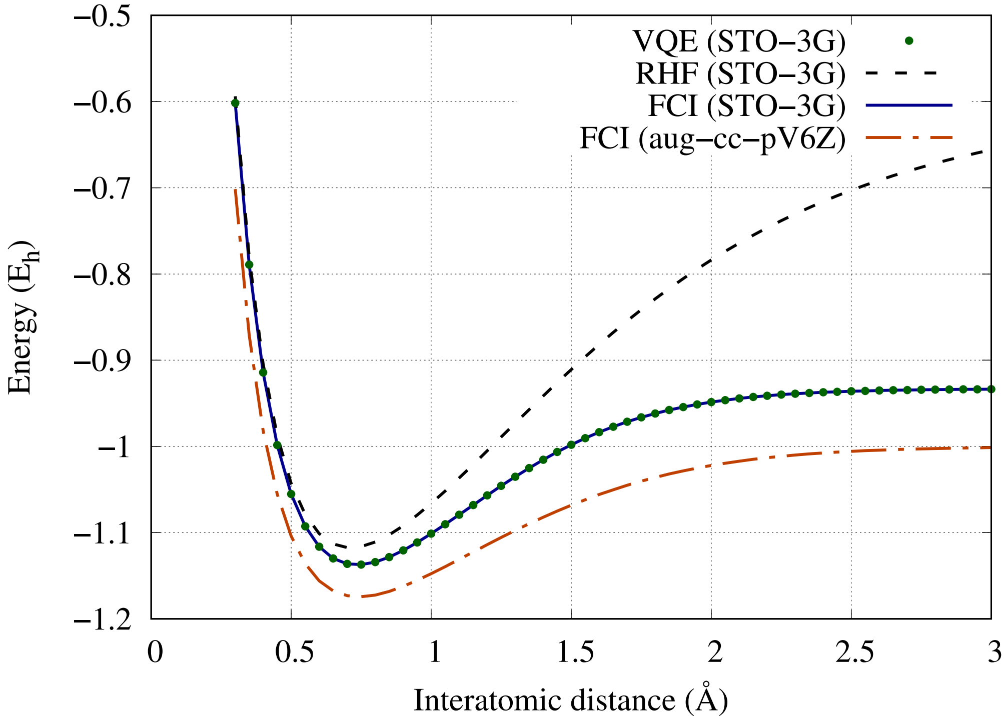

Figure 4–1: A plot showing potential energy curves for H_2 obtained using RHF, the VQE, and FCI.

energy at the equilibrium bond distance were employed in this case. Energy values were computed for interatomic distances ranging from 0.30 to 3.00 Å in increments of 0.05 Å. The resulting potential energy curves for all the tested methods are showed in Fig. 4–1. This graph shows that the potential energy curve for the reference state $|\Psi_{ref}\rangle$, which is obtained using the restricted Hartree-Fock method, remains fairly close to the FCI (STO-3G) curve for interatomic distances which are near or less than d_{eq}. For larger distances though, the reference state is unable to represent the wave function correctly using only a single Slater determinant. As for the VQE potential energy curve, it agrees once again with the FCI curve when the minimal STO-3G basis set is employed. However, it is not within chemical accuracy of the FCI curve which was obtained using the large aug-cc-pV6Z basis set. This further illustrates the importance of using a sufficiently large basis set when running quantum chemistry calculations.

4.3 Variable Directional Derivative Uncertainty Scheme

In order to obtain precise values for the total electronic energy when using a real quantum computer or Qiskit's QASM simulator, it is crucial to take an adequately large number of measurements during the final Hamiltonian averaging procedure. However, it is not necessary to take such a large number of measurements for each Hamiltonian averaging procedure in the VQE. Therefore, a scheme which varies the number of measurements for each iteration of the optimization algorithm was developed and implemented. The idea of using an adaptive strategy to improve the efficiency of the VQE was first mentionned by Hempel *et al.* [57].

Since the SPSA algorithm was utilized to optimize the trial state for this project, the variable uncertainty scheme works by decreasing the uncertainty in the directional derivative $\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)$ as its magnitude approaches zero. This requires setting up a relationship between the expected magnitude of the directional derivative $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp}$ and the uncertainty ε_d. A value for ε_d is needed to determine the number of measurements required according to eq. (3.93). When setting up the relationship, it is important to first specify minimum and maximum uncertainty values, $\varepsilon_{d,min}$ and $\varepsilon_{d,max}$, respectively. This prevents expectation values from being computed using too many measurements or too little.

In total, two different functions for selecting ε_d given $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp}$ were tested. The first one is a linear relationship, which is defined as:

$$\varepsilon_d = \begin{cases} \varepsilon_{d,max} & \text{if } |\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp} \geq A \\ \varepsilon_{d,min} & \text{if } |\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp} \leq B \\ c\left(|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp} - A\right) + \varepsilon_{d,max} & \text{otherwise} \end{cases}, \tag{4.2}$$

where A and B are user-defined parameters, and c corresponds to:

$$c = \frac{\varepsilon_{d,min} - \varepsilon_{d,max}}{B - A}. \tag{4.3}$$

The parameter A should be chosen such that it is near $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_0\right)|_{exp}$. If A is much larger than $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_0\right)|_{exp}$, part of the relationship for selecting ε_d between B and A could remain unused. This might lead to an increase in the total number of measurements taken. If A is much smaller than $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_0\right)|_{exp}$, the number of measurements taken could be too low as the optimization algorithm approaches the stationary point. This could make convergence difficult. As for B, it should be selected such that it is near the value $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|$ below which the optimization algorithm stops. If B is much smaller than that value, the final directional derivative might not be computed with the minimum uncertainty $\varepsilon_{d,min}$. This could result in a misidentification of the stationary point. On the other hand, employing a large value for B might unnecessarily increase the number of measurements taken near the stationary point.

The second function for selecting ε_d is an exponential relationship that can be written as:

$$\varepsilon_d = \begin{cases} \varepsilon_{d,max} & \text{if } |\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp} \geq A \\ \varepsilon_{d,min} & \text{if } |\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp} \leq B \\ a\left(\exp\left[-\lambda\left(|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp} - B\right)\right]\right) + b & \text{otherwise} \end{cases}, \tag{4.4}$$

where a and b correspond to:

$$a = \frac{\varepsilon_{d,min} - \varepsilon_{d,max}}{1 - \exp\left[-\lambda\left(A - B\right)\right]}, \tag{4.5}$$

and

$$b = \varepsilon_{d,min} - a. \tag{4.6}$$

Here, λ is a parameter that alters the shape of the exponential function. As a control, a constant function $\varepsilon_d = \varepsilon_{d,min}$ was also tested. Examples of uncertainty functions are shown in Fig. 4–2. In this figure, the effects of λ on exponential functions are apparent. For exponential functions with larger values of λ, the value for the uncertainty ε_d has a smaller rate of change near A, but a greater one near B.

In order to utilize the variable uncertainty scheme, the following steps need to be applied:

1. Select an uncertainty function. This requires choosing $\varepsilon_{d,min}$, $\varepsilon_{d,max}$, A, and B.
2. Set $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp}$ equal to A.
3. Use the uncertainty function to select a value for ε_d.
4. Calculate the number of measurements required $m_{d,req}$ by utilizing eq. (3.93) and ε_d.
5. Estimate $\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)$ and its magnitude using the Hamiltonian averaging procedure. Set the estimated magnitude equal to $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp}$.

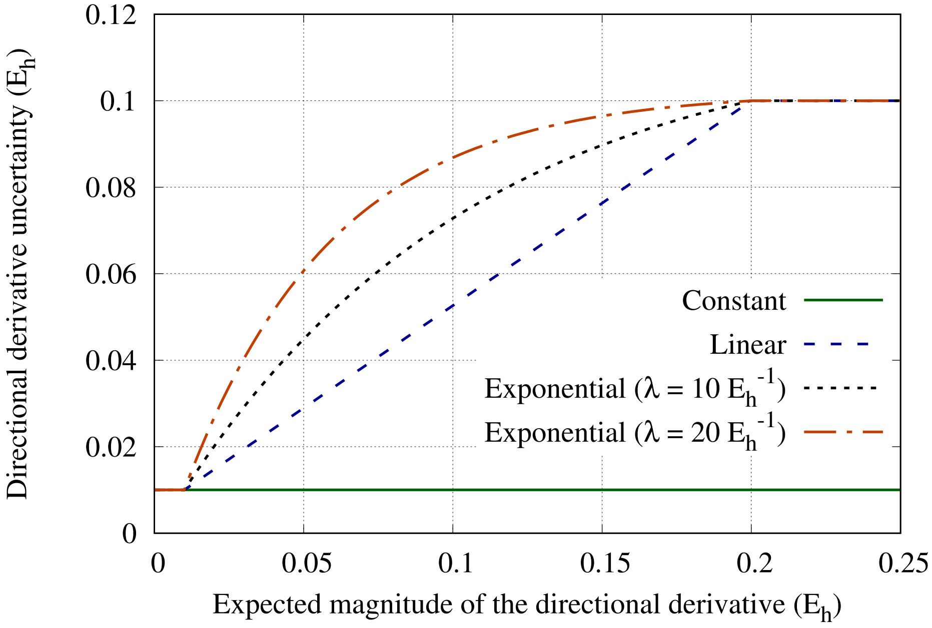

Figure 4–2: A plot containing functions used to select the uncertainty ε_d given the expected magnitude of the directional derivative $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp}$.

6. Use the uncertainty function to select a new value for the uncertainty $\varepsilon_{d,new}$. If $\varepsilon_{d,new}$ is greater than ε_d, proceed to Step 9. Otherwise, go to Step 7.
7. Calculate the new number of measurements required $m_{d,req,new}$ using eq. (3.93) and $\varepsilon_{d,new}$. If the inequality

   $$\Delta m_{d,req} = m_{d,req,new} - m_{d,req} > \Delta m_{d,req,min} \tag{4.7}$$

   is satisfied, go to Step 8. Otherwise, proceed to Step 9. Here, $\Delta m_{d,req,min}$ is the minimum increase in $m_{d,req}$ that is needed to run another Hamiltonian averaging procedure in the current iteration of the optimization algorithm. Furthermore, if $\Delta m_{d,req,min}$ is too small, computational resources may be used inefficiently since the estimate of the directional derivative only changes negligibly when $\Delta m_{d,req}$ is small.
8. Update the estimate of $\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)$ and its magnitude by taking $\Delta m_{d,req}$ additional measurements. Set ε_d equal to $\varepsilon_{d,new}$ and $|\widetilde{\nabla}_{\mathbf{v}} E\left(\boldsymbol{\theta}_k\right)|_{exp}$ equal to the magnitude of the new directional derivative estimate. Then, return to Step 6.
9. If the termination criteria of the optimization algorithm are satisfied, end the optimization procedure. Otherwise, go to the next iteration of the optimization procedure and set ε_d equal to $\varepsilon_{d,new}$. If $m_{d,req,new}$ was determined in the previous step, also set $m_{d,req}$ equal to $m_{d,req,new}$ and proceed to Step 5. If it was not, proceed to Step 4.

In order to test the variable uncertainty scheme, multiple VQE simulations were conducted for molecular hydrogen using Qiskit's QASM simulator. The goal of these simulations was to determine the average number of measurements required to reach the stationary point for different uncertainty functions. The equilibrium interatomic distance d_{eq} of 0.7414 Å was utilized for all simulations. The values of A and B were set to 0.2 and 0.01, respectively,

Table 4–2: The number of measurements required to reach the stationary point for H_2 at the equilibrium bond distance for various uncertainty functions

Uncertainty function	λ (E_h^{-1})	Number of measurements in millions			
		Average	Average lower bound	Average upper bound	Standard deviation
Constant	-	23.46	22.28	24.64	4.15
Linear	-	10.63	9.99	11.26	2.23
Exponential	5	10.10	9.47	10.72	2.20
Exponential	10	8.70	8.34	9.06	1.27
Exponential	20	8.41	8.01	8.81	1.40
Exponential	40	8.72	8.11	9.34	2.17

while $\varepsilon_{d,max}$ and $\varepsilon_{d,min}$ were set to 0.1 and 0.01, respectively. Additionally, z_d was set to 1 in eq. (3.93). As for the termination criterion of the optimization procedure, it was chosen such that the SPSA algorithm stopped once the estimate of $|\widetilde{\nabla}_{\mathbf{v}} E\left(\mathbf{t}_k\right)|$ was lower than 0.01. To put this into perspective, the difference between the expectation value $\langle \Psi_t|\hat{H}_e|\Psi_t\rangle$ for a directional derivative of 0.02 and the expectation value at the stationary point is less than 10^{-4} E_h when the UCCSD-based entangler is used for H_2 at d_{eq}. Results for these simulations are shown in Table 4–2. VQE simulations for each uncertainty function shown in the table were repeated 50 times. The columns labeled “Average lower bound” and “Average upper bound” correspond to the lower and upper bound of the 95% confidence interval associated with the average number of measurements.

Overall, Table 4–2 shows that varying the uncertainty of the directional derivative of energy is an effective way of reducing the number of measurements in the SPSA algorithm. For the particular case that was examined, using an exponential uncertainty function with a λ value of 20 was the optimal option. The uncertainty function needs to provide a balance between reducing the number of measurements at each iteration and increasing the total number of iterations. For the exponential uncertainty functions with lower values of λ, the number of measurements taken in the earlier iterations was unnecessarily large. Conversely, the function with a λ value of 40 computed directional derivatives that were too imprecise in the earlier iterations. This lead to an increase in the total number of iterations, which in turn lead to a rise in the total number of measurements.

It is important to note that further testing needs to be done in order to better quantify the efficiency of this variable uncertainty scheme. One aspect that needs to be studied is the effect of increasing the number of optimization parameters. For H_2 and the UCCSD-based ansatz, only the parameter associated with the double excitation operator is needed. Creating trial states $|\Psi_t\rangle$ for most molecules will require utilizing a larger number of parameters. Furthermore, all of the simulations shown in Table 4–2 were conducted using the same initial guess for the optimization parameter. The robustness of the scheme could be examined by randomly selecting initial values for $\boldsymbol{\theta}$ when testing various uncertainty functions. Another element that needs to be studied is the effect of the objective function's shape on the optimal choice of uncertainty function. However, the ultimate test would be to employ the scheme when running the VQE on noisy quantum computers.

CHAPTER 5
Implementation of Born-Oppenheimer Molecular Dynamics

5.1 Energy Gradient

As shown in Chapter 2, the goal of a Born-Oppenheimer molecular dynamics simulation is to predict the motion of the nuclei in a system of molecules by using the following equations:

$$M_I \ddot{\mathbf{R}}_I\left(t\right) = -\nabla_I \langle \Psi_{e,0} | \hat{H}_e | \Psi_{e,0} \rangle, \tag{5.1}$$

and

$$E_0 \Psi_{e,0} = \hat{H}_e \Psi_{e,0}. \tag{5.2}$$

In Chapters 3 and 4, it was demonstrated that an approximate solution to eq. (5.2) can be found in polynomial time by utilizing the VQE on a quantum computer. The outcome of a VQE simulation is the total electronic energy E, which corresponds to the expectation value $\langle \Psi_t | \hat{H}_e | \Psi_t \rangle$. Here, $|\Psi_t\rangle$ is an approximation to the exact ground state $|\Psi_{e,0}\rangle$.

In the case of eq. (5.1), its purpose is to predict how the position of a classically-behaving nucleus evolves in time. The first step in solving this equation is to obtain the gradient of the potential energy surface $\nabla_I \langle \Psi_t | \hat{H}_e | \Psi_t \rangle$ with respect to the coordinates of nucleus I. For conciseness, the gradient will also be written as $\nabla_I E\left(\mathbf{R}_j\right)$, where $\mathbf{R}_j$ contains the position of all nuclei for each iteration j. Every iteration is associated with a time

$$t_j = j \Delta t, \tag{5.3}$$

where Δt is the time step.

In order to estimate the gradient, it was decided to use the finite difference method [102]. With this approach, an approximation of a component of the gradient $\widetilde{\nabla}_{\hat{\mathbf{u}}} E\left(\mathbf{R}_j\right)$ is computed using the following equation:

$$\widetilde{\nabla}_{\hat{\mathbf{u}}} E\left(\mathbf{R}_j\right) = \frac{E\left(\mathbf{R}_j + \delta\hat{\mathbf{u}}\right) - E\left(\mathbf{R}_j - \delta\hat{\mathbf{u}}\right)}{2\delta}, \tag{5.4}$$

where δ is the step size and $\hat{\mathbf{u}}$ is a unit vector. The process of approximating the gradient is repeated for all N nuclei in the system. Overall, this requires computing the total electronic energy $6N$ times using the VQE.

It must also be emphasized that the selection of the step size plays a crucial role in the number of measurements required to approximate the gradient with a certain uncertainty. As for the directional derivative used in the classical optimization procedure, the number of measurements $m_{d,req}$ needed for each component of the gradient is determined using eq (3.93), which, as a reminder, corresponds to:

$$m_{d,req} = \left(\frac{\left(z_d\right)^2 M}{\delta^2 \varepsilon_d^2}\right) \sum_{i=1}^{M} h_i^2. \tag{5.5}$$

In this case, the step size must be sufficiently small to avoid introducing excessive numerical error in eq. (5.4), but it must also be sufficiently large to avoid taking too many measurements given an uncertainty ε_d. More details concerning eq. (5.5) are given in the Appendix.

5.2 Integration of the Equations of Motion

One of the last steps in every iteration of the BOMD method is to integrate the equations of motion given in eq. (5.1). For this research project, the velocity Verlet algorithm [103] was utilized to complete this task. Assuming that a particle has a single degree of freedom, the algorithm is able to determine its position R_j and its velocity $\dot{R}_j$ at iteration j. Knowing the initial position R_0 and the initial velocity $\dot{R}_0$ of a particle, R_j and $\dot{R}_j$ can be found by using the following equations:

$$R_j = R_{j-1} + \Delta t \dot{R}_{j-1} + \frac{(\Delta t)^2 f\left(R_{j-1}\right)}{2}, \tag{5.6}$$

and

$$\dot{R}_j = \dot{R}_{j-1} + \frac{\Delta t \left[f\left(R_j\right) + f\left(R_{j-1}\right)\right]}{2}. \tag{5.7}$$

In this case, $f\left(R\right)$ is found by rearranging the equations of motion such that it is written in this form:

$$\ddot{R} = f\left(R\right), \tag{5.8}$$

where $\ddot{R}$ is the acceleration of a particle. As a result, the position $x_{I,j}$ and the velocity $\dot{x}_{I,j}$ of nucleus I along the x-axis at time t_j is obtained by utilizing the equations:

$$x_{I,j} = x_{I,j-1} + \Delta t \dot{x}_{I,j-1} - \frac{(\Delta t)^2 \widetilde{\nabla}_{x_I} E\left(\mathbf{R}_{j-1}\right)}{2M_I}, \tag{5.9}$$

and

$$\dot{x}_{I,j} = \dot{x}_{I,j-1} - \frac{\Delta t \left[\widetilde{\nabla}_{x_I} E\left(\mathbf{R}_j\right) + \widetilde{\nabla}_{x_I} E\left(\mathbf{R}_{j-1}\right)\right]}{2M_I}. \tag{5.10}$$

It is important to note that the error associated with the nuclei positions when using the velocity Verlet algorithm is of order Δt^4, as demonstrated by Frenkel and Smit [104]. Therefore, care must be taken when selecting a time step. One way to verify that the chosen time step is sufficiently small is to compute the total energy of the system E_{sys} at each iteration. This energy corresponds to:

$$E_{sys} = \langle \Psi_t | \hat{H}_e | \Psi_t \rangle + \frac{1}{2} \sum_{I=0}^{N-1} M_I |\dot{\mathbf{R}}_I|^2, \tag{5.11}$$

where the first term is the sum of the potential energy in the system and the kinetic energy of the electrons, while the second term is the kinetic energy of the nuclei. If no external forces are applied to the system, E_{sys} should remain constant for the entirety of the simulation. Another measure than can be tracked is the velocity of the center of mass of the system. Its value should also remain constant during the simulation [104]. Therefore, if the total energy of the system or the velocity of the center of mass varies excessively during the simulation, the step size should be reduced.

CHAPTER 6
Dynamics of Molecular Hydrogen

6.1 Born-Oppenheimer Molecular Dynamics Simulation

In this section, results regarding simulations of the time evolution of an H_2 molecule will be presented. These simulations were conducted to ensure that the VQE could be used in the BOMD approach to find approximate solutions to the TISE given in eq. (5.2). They were inspired by those run by Wathelet *et al.* [105], who utilized Car-Parrinello molecular dynamics on a classical computer to detemine the equilibrium bond distance d_{eq} and the harmonic vibrational frequency ν of diatomic molecules. In order to mitigate the impact of anharmonicity on the results, Wathelet *et al.* used initial bond distances that were near d_{eq}. For this reason, the hydrogen nuclei were set 0.8 Å apart to begin the simulations of H_2. This is close to the experimental equilibrium bond distance of 0.7414 Å [97]. Additionally, the initial velocities were set to zero for both hydrogen nuclei. As for most of the VQE parameters, they were the same as those chosen for determining the potential energy surface of molecular hydrogen in Section 4.1.

In total, two simulations were conducted, one using the state vector simulator and another using the QASM simulator. In both cases, the time step in the velocity Verlet algorithm was set to 0.2 fs and the simulation was allowed to run for 14 fs. As for the numerical approximation to the energy gradient with respect to the nuclear coordinates, the step size was chosen as 0.001 Å for the state vector simulator, while it was chosen as 0.05 Å for the QASM simulator. The step size was increased for the QASM simulator in order to

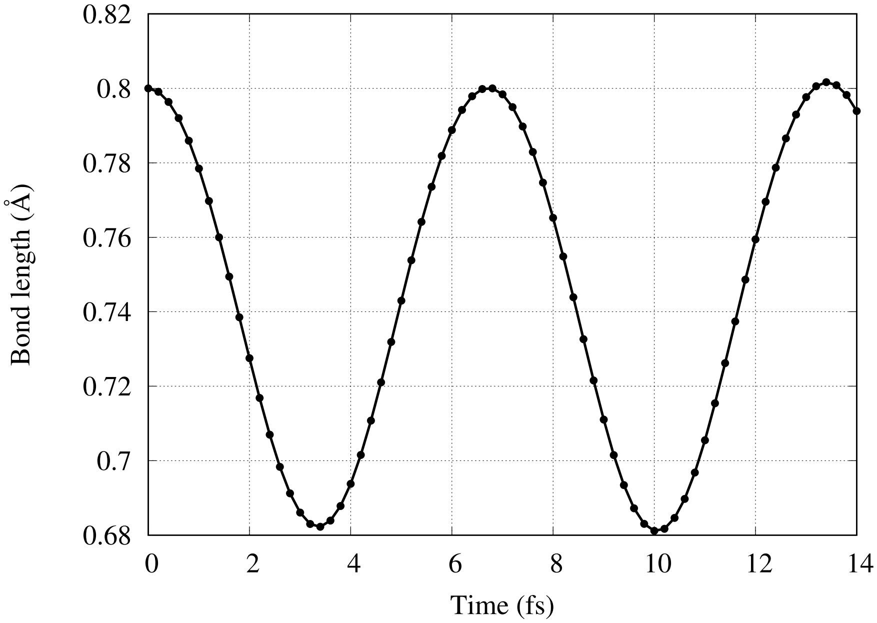

Figure 6–1: A plot showing the bond length of H_2 as a function of time. The results were obtained by a running a BOMD simulation on IBM'S QASM simulator.

reduce the number of measurements required to compute the gradient with an uncertainty of 0.005 E_h. To further accelerate the simulation, an exponential uncertainty function with a λ value of 20 and the same parameters as in Section 4.3 was utilized. It should also be noted that both the velocity Verlet algorithm and the numerical approximation to the energy gradient were implemented in C++ by the author of this thesis.

After running the simulations, the equilibrium bond length and the harmonic vibrational frequency were determined by examining the time evolution of the molecule's bond length. For the QASM simulation, a plot of the bond length as a function of time is presented in Fig. 6–1. To determine the equilibrium bond length, the mean of the minimum and maximum bond lengths in the first period was taken. As for the harmonic vibrational frequency ν, it

Table 6–1: A comparison of vibrational frequency and bond length results for H_2

Method	Harmonic vibrational frequency (cm^{-1})	Equilibrium bond length (Å)	QASM relative difference (%)	
			Harmonic vibrational frequency	Equilibrium bond length
BOMD (QASM)	4811	0.7408	-	-
BOMD (State vector)	4976	0.7388	3.32	0.275
FCI (STO-3G)	4953	0.7380	2.87	0.389
Experimental	4401.213	0.74144	9.31	0.080

is expressed as a wavenumber $\tilde{\nu}$. This wave number corresponds to [4]:

$$\tilde{\nu} = \frac{\nu}{c} = \frac{1}{Tc}, \tag{6.1}$$

where c is the speed of light and T is the period of oscillation.

To ensure that the values of d_{eq} and $\tilde{\nu}$ that were computed in both simulations are accurate, they were compared to results from experiments [97] and from a full configuration interaction simulation. The uncertainty associated with the experimental equilibrium bond length should be less than $\pm 9 \times 10^{-5}$Å while the one associated with the vibrational frequency might exceed ±0.010 cm^{-1} [97]. As for the FCI calculation, it was performed using Gaussian 16 [98]. The vibrational frequency was determined in Gaussian by computing the second derivative of energy with respect to the spatial coordinates of the nuclei at the stationary point. The STO-3G basis set was also used for this simulation. The results are summarized in Table 6–1. The last two columns in this table show the relative difference between the results obtained using the QASM simulator and those obtained using other methods.

For the vibrational frequency, there is a large error of 9.31% between the QASM value and the experimental one. This error is primarily due to the fact that the STO-3G basis set is not sufficiently large. In that basis set, the force on the nuclei is overestimated. Using the STO-3G basis set, the initial magnitude of the force on each nuclei was found to be 0.0442 Hartree per Bohr radius ($\mathrm{E_h/a_0}$), while for aug-cc-pV6Z, it was found to be 0.0340 $\mathrm{E_h/a_0}$. These values were obtained from FCI simulations conducted in Gaussian 16 [98]. In the presence of larger forces, the nuclei accelerate faster towards each other. This means that they have a shorter period of oscillation, and thus a larger vibrational frequency.

As for the difference between the frequencies calculated using the QASM simulator and the one obtained with FCI, it is mostly due to the uncertainty associated with measuring the total electronic energy and its gradient. Unfortunately, the number of measurements $m_{d,req}$ required to determine a component of the gradient with uncertainty ε_d is inversely proportional to ε_d^2 and δ^2. To decrease $m_{d,req}$, the gradient could be computed analytically instead. As a result, $m_{d,req}$ would no longer depend on δ^2. However, this approach requires deriving an analytical gradient equation for every entangler.

Another minor source of error is the size of the time step used in the velocity Verlet algorithm. The result is a deviation in the position and the velocity of the nuclei. To quantify the severity of the errors due to the uncertainty ε_d and the time step size, the total energy of the system E_{sys} can be tracked. In Fig. 6–2, a plot of the relative total energy drift for both BOMD simulations is shown. The relative total energy drift $\Delta_{rel}E_{sys}$ is defined as:

$$\Delta_{rel}E_{sys} = \frac{E_{sys} - E_{sys,0}}{E_{sys,0}}, \tag{6.2}$$

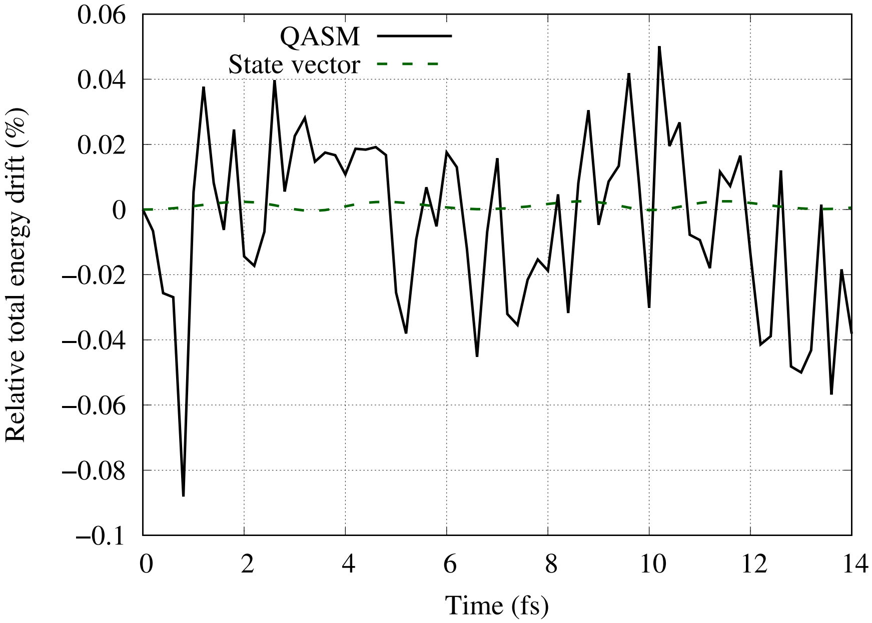

Figure 6–2: A plot showing the relative drift in total system energy for both the state vector and QASM simulations of H_2.

where $E_{sys,0}$ is the total system energy before the first time step. As expected, the total energy drift was less significant for the state vector simulation since ε_d is zero in that case.

For the equilibrium bond length, the sources of error are similar, but their impact is not as important. For instance, increasing the size of the basis set in this case does not result in a major change in d_{eq}. A source of error that could have slightly affected the results is anharmonicity. This would partly explain why the equilibrium bond length is larger in the BOMD simulations than in the FCI one. The anharmonicity causes the difference between the maximum bond length and d_{eq} to be larger than the one between d_{eq} and the minimum bond length. This is because the magnitude of the forces acting on the nuclei is greater at $d_{eq} - \Delta d$ than at $d_{eq} + \Delta d$, where Δd is a small change in bond length. Since d_{eq} was approximated in the BOMD simulations as the average of the minimum and maximum bond length, this lead to a slight increase in the estimated d_{eq}.

6.2 Optimization Parameter Extrapolation

There are a multitude of parameters that have an impact on the efficiency of the VQE. One of them is the initial guess for the optimization parameters $\boldsymbol{\theta}$. Since the VQE is run multiple times at each time point in a BOMD simulation, it is possible to utilize results found at previous time points to make an initial guess for $\boldsymbol{\theta}$. Because the same entangler is used throughout the simulation, the optimization parameters are not expected to vary greatly between time points if the time step size Δt is small. For a VQE simulation of H_2 with the UCCSD-based entangler and the STO-3G basis set, the relationship between the double excitation amplitude θ_{02}^{13} and the bond length is shown in Fig. 6–3.

In this section, three different approaches for selecting initial optimization parameters will be tested. The first one consists in taking advantage of perturbation theory by using

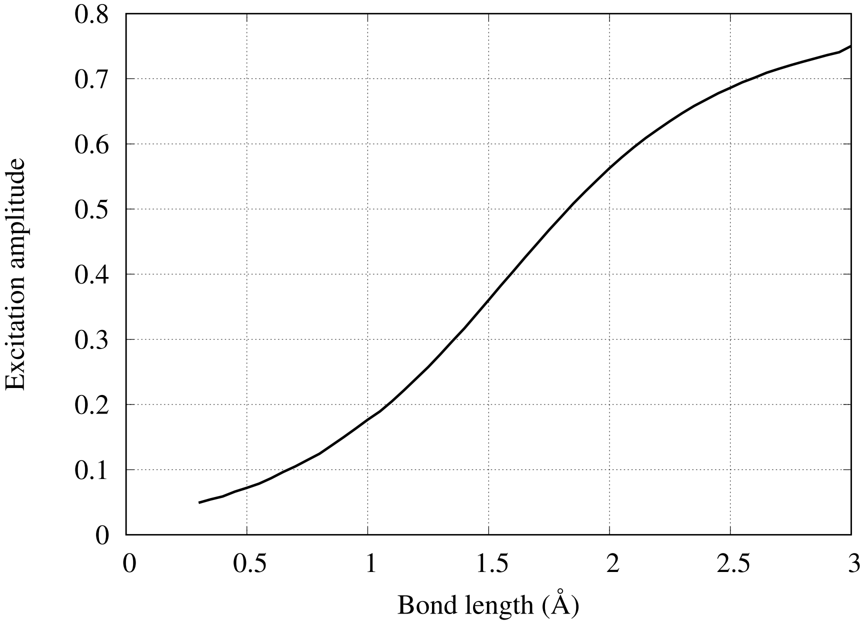

Figure 6–3: A plot showing the relationship between the double excitation amplitude in the UCCSD-based ansatz and the bond length.

eqs. (3.42) and (3.43) to calculate an initial guess. In the second approach, the optimal parameters from the previous time point are used to generate the first trial state. However, it is also possible to generalize this method by making use of results from multiple time points. Therefore, for the last approach, the optimal parameters from two previous time points are used to extrapolate a value for the new optimization parameters. To find the extrapolated value, the directional derivative of a parameter θ_i with respect to the spatial coordinates of the nuclei along a vector $\mathbf{v}$ is first approximated as:

$$\widetilde{\nabla}_{\mathbf{v}}\theta_i\left(\mathbf{R}_{s-1}\right) = \frac{\theta_{i,s-1} - \theta_{i,s-2}}{|\mathbf{v}|}. \tag{6.3}$$

Here, s represents the time point with the unknown optimization parameters and $\mathbf{v}$ is a vector from the nuclear coordinates at $s-2$ to those at $s-1$. Mathematically, $\mathbf{v}$ can be written as:

$$\mathbf{v} = \mathbf{R}_{s-1} - \mathbf{R}_{s-2}. \tag{6.4}$$

Similarly, another vector $\mathbf{w}$ is defined as:

$$\mathbf{w} = \mathbf{R}_{s} - \mathbf{R}_{s-2}. \tag{6.5}$$

The scalar projection p of $\mathbf{w}$ onto $\mathbf{v}$ is then determined using the following equation [106]:

$$p = \mathbf{w} \cdot \frac{\mathbf{v}}{|\mathbf{v}|}. \tag{6.6}$$

Afterwards, the optimization parameter $t_{i,s}$ is found by performing a linear extrapolation along $\mathbf{v}$ using the equation:

$$\theta_{i,s} = \theta_{i,s-2} + p\left[\widetilde{\nabla}_{\mathbf{v}}\theta_i\left(\mathbf{R}_{s-1}\right)\right]. \tag{6.7}$$

Table 6–2: A comparison of vibrational frequency and bond length results for H_2

Simulation	Zero initial nuclei velocity		
Method	Perturbation	Previous time point	Extrapolated
Average relative difference (%)	36.6	11.7	1.2
Maximum relative difference (%)	40.8	17.5	2.6

Simulation	Nonzero initial nuclei velocity (100 pm/fs)		
Method	Perturbation	Previous time point	Extrapolated
Average relative difference (%)	34.7	7.1	2.5
Maximum relative difference (%)	40.4	10.1	5.3

This procedure is then repeated for all the other optimization parameters.

To test these three approaches, BOMD simulations were conducted using the state vector simulator. Each approach was tested twice using a H_2 molecule, the UCCSD-based entangler, and the STO-3G basis set. In the first case, the hydrogen nuclei were placed 1.0 Å apart on the x-axis and given initial velocities of zero. The simulations were run for 14 fs using time steps of 0.5 fs. In the other case, the initial velocities were instead set to 100 pm/fs in the plus and minus y-direction. This was done to determine the impact of having nonparallel vectors **v** and **w**. The average and maximum relative difference between the optimal parameter θ_{02}^{13} and the initial guess was then calculated for each simulation. The results are shown in Table 6–2. It should be noted that the first two time points were not included in the averages since there are not enough previous results to use the extrapolated strategy before the third time point is reached.

From these results, it appears that using perturbation theory to calculate the initial optimization parameters is a poor approach. Using extrapolation to compute new parameters seems to be the best strategy. However, the accuracy of the extrapolated guesses deteriorated slightly when $\mathbf{v}$ and $\mathbf{w}$ were not parallel. The robustness of this method should be tested further. For example, additional simulations could be run using different entanglers and different molecules. The impacts of increasing the number of parameters or adding noise should also be studied. For large number of parameters, it might be advantageous to use the optimal parameters from the previous time point instead of spending computational power determining an extrapolated guess for every parameter. It should also be mentioned that the effectiveness of this method on reducing the number of optimization iterations depends greatly on the chosen optimization strategy.

CHAPTER 7
Born-Oppenheimer Molecular Dynamics using Superconducting Quantum Computers

7.1 Quantum Circuit Depth Reduction

The ultimate goal of this research project was to simulate the time evolution of H_2 using a real quantum computer. To achieve this, however, it was first necessary to reduce the depth of the quantum circuits used in a BOMD simulation. The depth of a circuit is defined as the number of time steps between the input state and the qubit measurements. Each time step can contain multiple quantum operations that are applied in parallel [107]. When the depth of a quantum circuit is too large, the qubits are more likely to be affected by interactions with the environment, resulting in decoherence. The outcome is a loss of accuracy [108]. Therefore, reducing the circuit depth is crucial since today's quantum computers are sensitive to environmental noise [48, 55].

In order to reduce the depth of a quantum circuit, one method is to utilize the Bravyi-Kitaev transformation [89, 109] instead of the Jordan-Wigner one to replace fermionic operators by Pauli operators. For a computer with m qubits, the former mapping procedure only generates $O\left(\log\left(m\right)\right)$ quantum gates per fermionic operator, while $O\left(m\right)$ gates are needed using the latter approach. When using the Bravyi-Kitaev transformation, the qubit state vector $\mathbf{b}_m$ is obtained from the fermionic occupation state vector $\mathbf{f}_m$ by using the following equation:

$$\mathbf{b}_m = \beta_m \mathbf{f}_m. \tag{7.1}$$

Here, β_m is a matrix which is defined as:

$$\beta_{2^x} = \left[\begin{array}{c|c} \beta_{2^{x-1}} & 0 \\ \hline 0 & \beta_{2^{x-1}} \\ \leftarrow 1 \rightarrow & \end{array}\right], \tag{7.2}$$

where $\leftarrow 1 \rightarrow$ indicates that the bottom row is filled with ones and β_1 corresponds to:

$$\beta_1 = \begin{bmatrix} 1 \end{bmatrix}. \tag{7.3}$$

For instance, the reference state for H_2 when using the STO-3G basis set is mapped from the fermionic occupation basis to the qubit basis in the following manner:

$$\mathbf{b}_4 = \beta_4 \mathbf{f}_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \tag{7.4}$$

where sums are computed using mod(2). As for the procedure utilized to generate the Pauli operators with the Bravyi-Kitaev transformation, it will not be presented here. The interested reader can refer to Ref. [89] where the procedure is described in great details. For the electronic Hamiltonian associated with the reference state shown in eq. (7.4), applying the transformation leads to this result:

$$\begin{aligned} \hat{H}_e = {} & h_0 + h_1 (Z_0 + Z_2) + h_2 (Z_0 Z_1 + Z_1 Z_2 Z_3) + h_3 (Z_1 + Z_1 Z_3) + h_4 Z_0 Z_2 \\ & + h_5 (X_0 X_2 Z_3 - X_0 Z_1 X_2 - X_0 Z_1 X_2 Z_3 + X_0 X_2) + h_6 (Z_0 Z_1 Z_2 Z_3 + Z_0 Z_1 Z_2) \\ & + h_7 Z_0 Z_2 Z_3, \end{aligned} \tag{7.5}$$

where h_i is a real coefficient.

In order to simplify the molecular Hamiltonian, a strategy used by O'Malley *et al.* [48] can be employed. They noted that only Z operators act on qubits 1 and 3. Since the maximum and minimum expectation values of Z occur when the qubit state is an eigenstate of Z, qubits 1 and 3 must be in states $|0\rangle$ or $|1\rangle$ to minimize $\langle \hat{H}_e \rangle$. Furthermore, because the reference state was obtained using the Hartree-Fock method, the configuration $|1110\rangle$ must be part of the optimal wave function. This means that qubits 1 and 3 are in states $|1\rangle$ and $|0\rangle$, respectively, in the optimal quantum state. As a result, the operators Z_1 and Z_3 can be replaced by their eigenvalues -1 and 1, respectively. After removing qubits 1 and 3, the electronic Hamiltonian becomes:

$$\hat{H}_e = h_0 - 2h_3 + (h_1 - h_2)\,(Z_0 + Z_1) + (h_4 - 2h_6 + h_7)\,Z_0 Z_1 + 4h_5 X_0 X_1, \tag{7.6}$$

where qubit 2 was relabeled qubit 1. Similarly, the reference state can be simplified to:

$$|\Psi_{ref}\rangle = |11\rangle. \tag{7.7}$$

Since two qubits were removed, it is also necessary to simplify the UCCSD-based ansatz. This can be achieved by following a strategy that was developed by Hempel *et al.* [57]. The first step is to recognize that the single excitation operators have no effect on the reference state. This is due to the Brillouin theorem which states that [12]:

$$\langle \Psi_{HF} | \hat{H}_e a_a^\dagger a_i | \Psi_{HF} \rangle = 0, \tag{7.8}$$

where $|\Psi_{HF}\rangle$ is a Hartree-Fock state and $a_a^\dagger a_i|\Psi_{HF}\rangle$ is an excited state generated by interchanging a virtual orbital and an occupied one. Additionally, since

$$\langle\Psi_{HF}|\hat{H}|\Psi_{HF}\rangle \leq \langle\Psi_{HF}|a_i^\dagger a_a \hat{H}_e a_a^\dagger a_i|\Psi_{HF}\rangle, \tag{7.9}$$

the presence of a singly-excited Hartree-Fock state can only increase the expectation value of the electronic Hamiltonian. Therefore, the single excitation operators can be removed and the ansatz becomes:

$$\begin{aligned} U = &\exp\left(-\frac{i\theta_{0,2}^{1,3}}{8}Y_0X_2\right) \times \exp\left(-\frac{i\theta_{0,2}^{1,3}}{8}X_0Y_2\right) \times \exp\left(\frac{i\theta_{0,2}^{1,3}}{8}Y_0Z_1X_2Z_3\right) \\ &\times \exp\left(\frac{i\theta_{0,2}^{1,3}}{8}X_0Z_1Y_2Z_3\right) \times \exp\left(\frac{i\theta_{0,2}^{1,3}}{8}X_0Z_1Y_2\right) \times \exp\left(\frac{i\theta_{0,2}^{1,3}}{8}Y_0Z_1X_2\right) \\ &\times \exp\left(-\frac{i\theta_{0,2}^{1,3}}{8}X_0Y_2Z_3\right) \times \exp\left(-\frac{i\theta_{0,2}^{1,3}}{8}Y_0X_2Z_3\right). \end{aligned} \tag{7.10}$$

Once again, the Z operators which are applied to qubits 1 and 3 of the unsimplified reference state can be replaced by their eigenvalues. The result is the following:

$$U = \exp\left(-\frac{i\theta_{0,2}^{1,3}}{2}Y_0X_1\right)\exp\left(-\frac{i\theta_{0,2}^{1,3}}{2}X_0Y_1\right), \tag{7.11}$$

where the Pauli operator indices were relabeled to match those of the simplified Hamiltonian. Because the reference state contains a single configuration, it is also possible to combine the two remaining exponential operators into one [57]. This is achieved by using the fact that a Pauli operator can be written as the product of the other two Pauli operators. For instance, X can be expressed as $-iYZ$, while Y can be expressed as iXZ. Thus, the ansatz can be written as:

$$U = \exp\left(-\frac{i\theta_{0,2}^{1,3}}{2}Y_0X_1\right)\exp\left(-\frac{i\theta_{0,2}^{1,3}}{2}Y_0Z_0X_1Z_1\right). \tag{7.12}$$

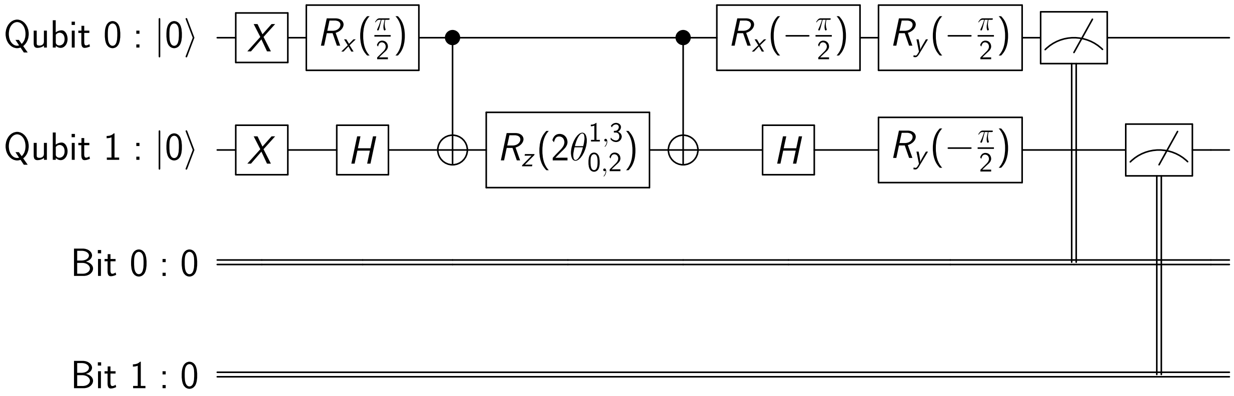

Figure 7–1: The largest quantum circuit obtained after simplifying the UCCSD-based ansatz and the electronic Hamiltonian.

Replacing the Z operators by their eigenvalues, the ansatz becomes:

$$U = \exp\left(-i\theta_{0,2}^{1,3} Y_0 X_1\right). \tag{7.13}$$

After these simplifications, the largest quantum circuit that needs to be submitted to the quantum computer is the one which requires computing the expectation value $\langle X_0 X_1 \rangle$. This circuit is shown in Fig. 7–1. It has a circuit depth of 7 gates. By comparison, the largest quantum circuit that was used in the previous sections of this thesis had a circuit depth of 73 gates. However, it is possible to create an ansatz for H_2 in the STO-3G basis set that has an even smaller circuit depth. This will be shown in the following section.

7.2 Minimal Circuit Depth Ansatz

To maximize the chance of successfully conducting a BOMD simulation on a quantum computer, it was decided to create an ansatz which is specific to the molecule and the basis set of interest. As a starting point, it is helpful to determine the form of the possible trial states. For H_2 in the STO-3G basis set, the trial states all have the following form:

$$|\Psi_t\rangle = a\left(\theta_{0,2}^{1,3}\right)|1010\rangle + b\left(\theta_{0,2}^{1,3}\right)|0101\rangle, \tag{7.14}$$

where $a\left(\theta_{0,2}^{1,3}\right)$ and $b\left(\theta_{0,2}^{1,3}\right)$ are real coefficients, while $|1010\rangle$ and $|0101\rangle$ are fermionic occupation state vectors. Applying the Bravyi-Kitaev transformation, the trial states then become:

$$|\Psi_t\rangle = a\left(\theta_{0,2}^{1,3}\right)|1110\rangle + b\left(\theta_{0,2}^{1,3}\right)|0100\rangle. \tag{7.15}$$

Qubits 1 and 3 can afterwards be removed as was done in Section 7.1. This results in the following trial states:

$$|\Psi_t\rangle = a\left(\theta_{0,2}^{1,3}\right)|11\rangle + b\left(\theta_{0,2}^{1,3}\right)|00\rangle. \tag{7.16}$$

Previously, $|11\rangle$ was used as the reference state. However, initializing this state requires applying an X gate to each qubit. To minimize the circuit depth, the reference state is instead chosen to be $|00\rangle$. The goal of the entangler is therefore to transform state $|00\rangle$ into one of the trial states, as shown in this equation:

$$U\left(\theta_{0,2}^{1,3}\right)|00\rangle = a\left(\theta_{0,2}^{1,3}\right)|11\rangle + b\left(\theta_{0,2}^{1,3}\right)|00\rangle. \tag{7.17}$$

It turns out that it is possible to generate the trial states by using an ansatz that only contains two quantum gates. This ansatz corresponds to:

$$U\left(\gamma\right) = U_{CN,0,1}R_{y0}\left(\gamma\right), \tag{7.18}$$

where γ can be related to $\theta_{0,2}^{1,3}$ by noting that

$$\exp\left(-i\theta_{0,2}^{1,3}Y_0X_1\right)|11\rangle = U_{CN,0,1}R_{y0}\left(\gamma\right)|00\rangle. \tag{7.19}$$

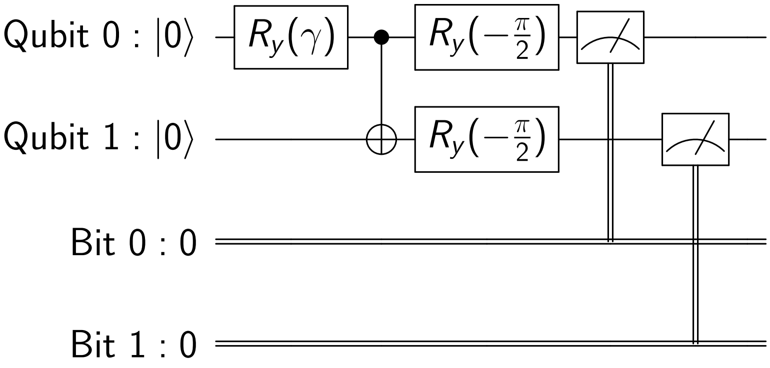

Figure 7–2: The largest quantum circuit obtained by using the ansatz given in eq. (7.18).

The parameter γ can then be expressed in terms of $\theta_{0,2}^{1,3}$ by solving the following equations:

$$\sin\left(\theta_{0,2}^{1,3}\right) = -\cos\left(\frac{\gamma}{2}\right), \tag{7.20}$$

and

$$\cos\left(\theta_{0,2}^{1,3}\right) = \sin\left(\frac{\gamma}{2}\right). \tag{7.21}$$

Using the simplified electronic Hamiltonian, which was derived in Section 7.1, the depth of the largest quantum circuit is reduced to only 3 gates. This circuit is shown in Figure 7–2. The smallest circuit, on the other hand, has a depth of 2 gates. It is needed for terms in the electronic Hamiltonian that only involve Z operators.

To verify if the ansatz given in eq. (7.18) can be used to run a BOMD simulation using one of IBM's superconducting quantum computers, calculations to determine the expectation value $\langle Z_0 Z_1 \rangle$ for a certain parameter γ were performed. In this case, γ was set to -2.789, which corresponds to the optimal parameter for an H_2 molecule with an interatomic distance of 1.0 Å. The expectation value was approximated using three different quantum computers. One was located in Melbourne and had 14 qubits, while the other two were in Tenerife and

Table 7–1: A comparison of the measurement probabilities and the expectation value $\langle Z_0 Z_1 \rangle$ calculated using different quantum computers

Quantum computer	Melbourne (14 qubits)	Tenerife (5 qubits)	Yorktown (5 qubits)	State vector simulator
$p_{\|00\rangle}$ (%)	5.15	10.88	1.91	3.08
$p_{\|01\rangle}$ (%)	6.05	11.57	7.02	0.00
$p_{\|10\rangle}$ (%)	6.05	22.95	4.38	0.00
$p_{\|11\rangle}$ (%)	82.75	54.60	86.68	96.92
$\langle Z_0 Z_1 \rangle$	0.758	0.310	0.772	1.000

Yorktown, but had 5 qubits. For each quantum computer, the 2 qubits with the lowest single-qubit and two-qubit gate error rates were selected. To verify the accuracy of these computations, the expectation value was also determined using the state vector simulator. The values that were obtained for $\langle Z_0 Z_1 \rangle$ are given in Table 7–1 along with the probability associated with each measurement outcome.

An important task that had to be completed before comparing the results presented in this table was to determine the uncertainty ε of the Hamiltonian averaging procedure. For the calculations performed on quantum computers, 24 576 measurement were taken. To compute the uncertainty, eq. (8.14), which is explained in further details in the Appendix, was employed. Since the variance of $\langle Z_0 Z_1 \rangle$ is 1, the equation simplifies to:

$$\varepsilon^2 = \frac{z^2}{m}. \tag{7.22}$$

For a z value of 3, the uncertainty turns out to be 0.019. In other words, there is a 99.7% probability of calculating an expectation value that is within 0.019 of the exact value of 1.000 when taking 24 576 measurements. This demonstrates that the inaccuracy of the results which were obtained using quantum computers is due to systematic errors. A common

Table 7–2: A comparison between expectation values computed by the IBM quantum computer in Yorktown and those found using the state vector simulator

Quantum computer	$\langle Z_0 \rangle$	$\langle Z_1 \rangle$	$\langle Z_0 Z_1 \rangle$	$\langle X_0 X_1 \rangle$	$\langle \hat{H}_e \rangle$ (E_h)
Yorktown (5 qubits)	-0.8213	-0.8741	0.7719	-0.1258	-1.0114
State vector simulator	-0.9384	-0.9384	1.0000	-0.3464	-1.1012

source of such errors is crosstalk [110]. In a superconducting quantum computer, the qubits are electromagnetically coupled to transmission lines such that they can be controlled and entangled. However, these electromagnetic interactions can also affect qubits in an undesirable way. For instance, applying an $R_y(\gamma)$ gate to one qubit could affect the state of another qubit. Apart from crosstalk, errors can also be caused by sending imperfect microwave pulse sequences to the qubits [48]. As an example, an imprecise pulse could lead to the application of a gate $R_y(\gamma + \delta_\gamma)$ instead of the intended $R_y(\gamma)$, where δ_γ is the deviation from the correct angle γ. It is also common for errors to affect the state preparation and the measurement procedures [111].

To further demonstrate the impact of these errors, the total electronic energy was determined for H_2 with an interatomic distance of 1.0 Å. This was accomplished by taking the same number of measurements as for Table 7–1 for each quantum circuit. The energy was computed using the 5-qubit quantum computer located in Yorktown and the state vector simulator. The expectation values associated with each term of the electronic Hamiltonian and the total electronic energy are given in Table 7–2 for both methods. Additionally, the energy determined using the quantum computer in Yorktown was compared to values obtained using the restricted Hartree-Fock (RHF) method and full configuration interaction

Table 7–3: The absolute error in total electronic energy for the IBM quantum computer in Yorktown and methods relying exclusively on classical computers

Method	RHF	Yorktown (5 qubits)	State vector simulator	FCI	FCI
Basis set	STO-3G	STO-3G	STO-3G	STO-3G	aug-cc-pV6Z
Total electronic energy (E_h)	-1.0661	-1.0114	-1.1012	-1.1012	-1.1476
Absolute energy error with respect to STO-3G (mE_h)	35.1	89.8	0.0	0.0	-
Absolute energy error with respect to aug-cc-pV6Z (mE_h)	81.5	136.2	46.4	46.4	0.0

(FCI). The FCI calculations were performed in Gaussian 16 [98] using both the STO-3G and aug-cc-pV6Z basis sets. The results are presented in Table 7–3 along with the absolute energy error.

The magnitude of the error values given in Table 7–3 demonstrates that it is not possible to conduct a BOMD simulation with one of IBM's quantum computers by measuring $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle$ at the theoretically optimal VQE parameters $\boldsymbol{\theta}$. Furthermore, chemical accuracy (1.6 mE_h) was not reached using the quantum computer in Yorktown. In fact, the quantum computer provided a worse energy approximation than the restricted Hartree-Fock method, which was employed to contruct the reference state. To obtain better results, an error-mitigation scheme could be utilized. So far, methods such as Richardson extrapolation [47, 112] and probabilistic error cancellation [112] have been proposed to reduce the impact of errors. As an alternative, a hardware-based entangler, such as the one created by

Kandala *et al.* [55], could be used to find the experimentally optimal VQE parameters that minimize $\langle \Psi_t | \hat{H}_e | \Psi_t \rangle$. Since the VQE takes advantage of the variation principle, the energy computed at the experimentally optimal parameters is either equal or more accurate than the energy found at the theoretically optimal parameters [53].

CHAPTER 8
Conclusion

8.1 Concluding Remarks

The advent of increasingly powerful quantum computers promises to revolutionize fields such as computational chemistry. It will become possible to solve problems that are intractable on classical computers. Such problems can be found in areas like combustion, materials science, and drug design. Therefore, the goal of the present work was to develop numerical schemes and methods to perform BOMD simulations on a quantum computer. These schemes were tested on a H_2 molecule using a quantum computer simulator. The results showed that it was theoretically possible to use the VQE to conduct a BOMD simulation on a quantum computer. However, it should be noted that results which are within chemical accuracy cannot be obtained using the minimal basis set that was utilized throughout this thesis. A larger basis set is required to attain that level of accuracy. This implies running simulations with a larger number of quantum gates and more qubits.

Furthermore, it was shown that techniques such as linear extrapolation could be employed to improve the initial guess of the VQE's optimization parameters. For simulations involving H_2, the average relative difference between the initial guesses and the optimal parameters was improved from 35%, using perturbation theory, to 2.5%, using linear extrapolation. This technique can also be applied to other entanglers than the UCCSD-based ansatz which was utilized in this thesis. However, further tests still need to be completed using more complex molecules and a larger number of parameters in order to better quantify the effectiveness of this strategy.

Additionally, the efficiency of the VQE was improved by allowing the number of measurements taken to vary for each iteration of the classical optimization procedure. That was accomplished by establishing relationships, known as uncertainty functions, between the uncertainty of the measured directional derivative of energy and the expected value of that derivative. The directional derivative in this case is taken with respect to the optimization parameters. Using an exponential uncertainty function, the number of measurements needed to determine the total electronic energy of H_2 was reduced by a factor of almost 3. Although this strategy was used for the SPSA algorithm, it could also be applied to other gradient-based algorithms.

8.2 Future Work

In order to reach the goal of running a molecular dynamics simulation using a quantum computer, some modifications need to be made to the VQE algorithm that was implemented for this thesis. One important aspect is the need to perform error mitigation. A strategy similar to the one developed concurrently by Li and Benjamin [47] as well as Temme *et al.* [112] could be employed. Their approach works by intentionally increasing the noise on the quantum computer and then extrapolating to the zero-noise limit to obtain more accurate results. Another modification that could be made is to use a hardware-based entangler instead of a chemically-inspired one like the UCCSD-based ansatz. Such an entangler was employed by Kandala *et al.* [55] to determine the PES of H_2, LiH, and BeH_2 using superconducting qubits. This type of entangler generates trial states by taking advantage of the natural entanglement present on a quantum computer. It is therefore less reliant on the device's ability to apply quantum gates accurately.

An additional aspect of the BOMD method that must be improved is the calculation of the energy gradient with respect to the spatial coordinates of the nuclei. The finite difference approach that was used in this thesis has several drawbacks. One of them is that the number of energy evaluations needed to compute the gradient is $6N$, where N is the number of nuclei. Furthermore, the number of quantum measurements required to achieve a target precision ε_d is proportional to $1/\delta^2$, where δ is the step size employed in the finite difference method. An alternative approach is to compute the gradient analytically. This method has the advantage of being less computationally expensive than its numerical counterpart [6]. Recently, analytical expressions for the energy gradient were determined by O'Brien *et al.* [113] and Mitarai *et al.* [114] for the VQE. A third strategy that is also worth considering is automatic differentiation, proposed by Tamayo-Mendoza *et al.* [115], which calculates the gradient by analytically evaluating the derivative of every operation in an algorithm. The main advantage of this differentiation technique is that it can compute gradients with machine precision without the need to derive analytical expressions for every algorithm.

Finally, it is also important to improve the variable uncertainty scheme which is employed to reduce the number of measurements taken in the VQE. At the moment, the efficiency of this method depends greatly on the several parameters which need to be selected for the uncertainty functions. As an alternative, the parameters for the uncertainty function could be optimized automatically during the optimization algorithm of the VQE. This could be achieved, for example, by computing multiple estimates of the directional derivative of energy for a single iteration. A different uncertainty value would be used to obtain each estimate. The energy at the next iteration would then be calculated for each directional derivative. Afterwards, the results would be compared to update the parameters of the uncertainty function. The parameters would only be updated at every few iteration. As a consequence, the variable uncertainty scheme would be more robust and would not depend as much on the initially selected parameters for the uncertainty function.

Appendix: Hamiltonian Averaging Uncertainty

Being able to predict the uncertainty of the Hamiltonian averaging procedure is crucial for obtaining meaningful results when running the VQE. For this reason, equations which approximate the number of measurements required to evaluate the expectation value $\langle \Psi_t | \hat{H}_e | \Psi_t \rangle$ and its directional derivative with a certain amount of uncertainty will be derived in this appendix. The approach employed to do so in this section is similar to those used by Wecker *et al.* [93] and McClean *et al.* [54].

For the first step, the electronic Hamiltonian can be reexpressed as:

$$\langle \Psi_t | \hat{H}_e | \Psi_t \rangle = h_0 + \sum_{i=1}^{M} \langle \Psi_t | \hat{H}_{e,i} | \Psi_t \rangle = h_0 + \sum_{i=1}^{M} h_i \langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle, \tag{8.1}$$

where h_0 and h_i are constants, $\hat{H}_{e,i}$ is a nonconstant term in $\hat{H}_e$, and M is the number of nonconstant terms in $\hat{H}_e$. Additionally, $\hat{P}_{z,i}$ corresponds to a product of Pauli-Z matrices while $|\xi_t\rangle$ is the state $|\Psi_t\rangle$ after changing the basis of the qubits in order to take a measurement. In a similar manner, the value which is calculated during the Hamiltonian averaging procedure, known as the estimator $\langle \widehat{\Psi_t | \hat{H}_e | \Psi_t} \rangle$, is defined as:

$$\langle \widehat{\Psi_t | \hat{H}_e | \Psi_t} \rangle = h_0 + \sum_{i=1}^{M} \langle \widehat{\Psi_t | \hat{H}_{e,i} | \Psi_t} \rangle = h_0 + \sum_{i=1}^{M} h_i \langle \widehat{\xi_t | \hat{P}_{z,i} | \xi_t} \rangle, \tag{8.2}$$

where $\langle \widehat{\Psi_t | \hat{H}_{e,i} | \Psi_t} \rangle$ and $\langle \widehat{\xi_t | \hat{P}_{z,i} | \xi_t} \rangle$ are the estimator of their respective expectation value. Each estimator $\langle \widehat{\Psi_t | \hat{H}_{e,i} | \Psi_t} \rangle$ corresponds to the average value $\bar{X}_{m_i}$ associated with the independent random variables X_k, where m_i is the number of measurements taken.

Afterwards, the central limit theorem is applied. The theorem states that if m_i is sufficiently large, the distribution of $\bar{X}_{m_i}$ for a random variable X_k with mean μ_i and variance σ_i^2 is approximately normal. The expectation value $\mathrm{E}\left[\bar{X}_{m_i}\right]$ and the variance $\mathrm{Var}\left[\bar{X}_{m_i}\right]$ of this normal distribution are defined as follows [116]:

$$\mathrm{E}\left[\bar{X}_{m_i}\right] = \mu_i, \tag{8.3}$$

$$\mathrm{Var}\left[\bar{X}_{m_i}\right] = \frac{\sigma_i^2}{m_i}. \tag{8.4}$$

As a result, the variance of the estimator $\widehat{\langle \Psi_t | \hat{H}_{e,i} | \Psi_t \rangle}$ corresponds to:

$$\mathrm{Var}\left[\widehat{\langle \Psi_t | \hat{H}_{e,i} | \Psi_t \rangle}\right] = \frac{\mathrm{Var}\left[\langle \Psi_t | \hat{H}_{e,i} | \Psi_t \rangle\right]}{m_i}. \tag{8.5}$$

With this equation, it is then possible to determine the variance of the estimator $\widehat{\langle \Psi_t | \hat{H}_e | \Psi_t \rangle}$. To do so, the variance of the sum of two random variables $\bar{X}_{m_i}$ and $\bar{Y}_{m_j}$ needs to be obtained. This can be done by using the following equation [117]:

$$\mathrm{Var}\left[\bar{X}_{m_i} + \bar{Y}_{m_j}\right] = \mathrm{Var}\left[\bar{X}_{m_i}\right] + \mathrm{Var}\left[\bar{Y}_{m_j}\right] + 2\mathrm{Cov}\left[\bar{X}_{m_i}, \bar{Y}_{m_j}\right], \tag{8.6}$$

where $\mathrm{Cov}\left[\bar{X}_{m_i}, \bar{Y}_{m_j}\right]$ is the covariance between $\bar{X}_{m_i}$ and $\bar{Y}_{m_j}$. Since all estimators $\widehat{\langle \Psi_t | \hat{H}_{e,i} | \Psi_t \rangle}$ are independent of each other, the covariance will always be zero. Therefore, the variance of $\widehat{\langle \Psi_t | \hat{H}_e | \Psi_t \rangle}$ corresponds to:

$$\mathrm{Var}\left[\widehat{\langle \Psi_t | \hat{H}_e | \Psi_t \rangle}\right] = \sum_{i=1}^{M} \mathrm{Var}\left[\widehat{\langle \Psi_t | \hat{H}_{e,i} | \Psi_t \rangle}\right]. \tag{8.7}$$

It is also important to note that the sum of two independent random variables that are normally distributed follows a normal distribution [118]. For this reason, the uncertainty ε

is defined according to this equation:

$$\varepsilon = z\sqrt{\mathrm{Var}\left[\widehat{\langle\Psi_t|\hat{H}_e|\Psi_t\rangle}\right]}, \tag{8.8}$$

where

$$z = \frac{\bar{X}_m - \mathrm{E}\left[\bar{X}_m\right]}{\sqrt{\mathrm{Var}\left[\bar{X}_m\right]}} = \frac{\langle\Psi_t|\hat{H}_e|\Psi_t\rangle - \widehat{\langle\Psi_t|\hat{H}_e|\Psi_t\rangle}}{\sqrt{\mathrm{Var}\left[\widehat{\langle\Psi_t|\hat{H}_e|\Psi_t\rangle}\right]}}. \tag{8.9}$$

The variable z corresponds to the z score associated with the upper limit of the interval $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle \pm \varepsilon$ [118]. If z is set to 1 for example, this means that in 68% of averaging procedures, $\widehat{\langle\Psi_t|\hat{H}_e|\Psi_t\rangle}$ is at most ε away from $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle$.

Next, in order to express the uncertainty in terms of the number of measurements taken for each estimator $\widehat{\langle\Psi_t|\hat{H}_{e,i}|\Psi_t\rangle}$, eqs. (8.8) and (8.5) are substituted into eq. (8.7). The result is shown in this equation:

$$\varepsilon^2 = z^2 \sum_{i=1}^{M} \left(\frac{\mathrm{Var}\left[\langle\Psi_t|\hat{H}_{e,i}|\Psi_t\rangle\right]}{m_i} \right). \tag{8.10}$$

It is then possible to rewrite eq. (8.10) in terms of the variance of $\langle\xi_t|\hat{P}_{z,i}|\xi_t\rangle$ by using the following rule [117]:

$$\mathrm{Var}\left[c\bar{X}_m\right] = c^2\mathrm{Var}\left[\bar{X}_m\right], \tag{8.11}$$

where c is a constant. The outcome therefore corresponds to:

$$\varepsilon^2 = z^2 \sum_{i=1}^{M} \left(\frac{h_i^2\mathrm{Var}\left[\langle\xi_t|\hat{P}_{z,i}|\xi_t\rangle\right]}{m_i} \right). \tag{8.12}$$

For the next step of this derivation, the uncertainty must be expressed in terms of the total number of measurements m taken during the Hamiltonian averaging procedure. The most efficient way to do so is to set the number of measurements m_i such that they are proportional to $|h_i|$ [93]. However, this approach was not chosen for this research project. This is because each job that is sent to one of IBM's quantum computer using Qiskit [91] contains an array of quantum circuits. Unfortunately, the number of measurements must be the same for all circuits in the array. The alternative would be to group circuits with the same m_i together and to submit a job for each of these groups. Since each job that is submitted is placed in a queue before it is processed though, this would drastically increase the time required to run a simulation. For this reason, it was decided to assign the same number of measurements to all circuits. Mathematically, this corresponds to:

$$m_i = \frac{m}{M}. \tag{8.13}$$

Substituting this equation into eq. (8.12) then gives the following:

$$\varepsilon^2 = \left(\frac{Mz^2}{m}\right) \sum_{i=1}^{M} h_i^2 \mathrm{Var}\left[\langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle\right]. \tag{8.14}$$

The only variable that is unknown in eq. (8.14) is the variance of $\langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle$. Fortunately, there is an upper bound to this variance. This maximum value occurs when $\langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle$ is zero. The only possible measurement outcomes in this case are 1 or -1. This results in a variance of 1. Therefore, the variance of $\langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle$ can be expressed as:

$$\mathrm{Var}\left[\langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle\right] \leq 1. \tag{8.15}$$

Substituting this inequality into eq. (8.14) then gives:

$$m \leq \left(\frac{Mz^2}{\varepsilon^2}\right) \sum_{i=1}^{M} h_i^2. \tag{8.16}$$

Assuming that z is set to 1, choosing the upper bound of this equation for m means that $\widehat{\langle\Psi_t|\hat{H}_e|\Psi_t\rangle}$ will be at most ε away from $\langle\Psi_t|\hat{H}_e|\Psi_t\rangle$ in at least 68% of averaging procedures. This statement is true as long as m is sufficiently large.

For the energy derivative with respect to a variable x, a similar approach to the one given above will be used to determine the number of measurements required to obtain a certain uncertainty ε_d. In this case, the energy $E\left(x\right)$ is defined as the expectation value of the electronic Hamiltonian and the derivative is computed using the finite difference method [102]. The numerical approximation of the energy derivative as well as the associated estimator are respectively given in the two following equations:

$$\frac{\mathrm{d}E\left(x\right)}{\mathrm{d}x} \approx \left(\frac{\mathrm{d}E\left(x\right)}{\mathrm{d}x}\right)_{num} = \frac{E\left(x+\delta\right) - E\left(x-\delta\right)}{2\delta}, \tag{8.17}$$

$$\frac{\widehat{\mathrm{d}E\left(x\right)}}{\mathrm{d}x} \approx \left(\frac{\widehat{\mathrm{d}E\left(x\right)}}{\mathrm{d}x}\right)_{num} = \frac{\widehat{E\left(x+\delta\right)} - \widehat{E\left(x-\delta\right)}}{2\delta}, \tag{8.18}$$

where δ corresponds to the step size. Since the energy estimators are independent random variables, the variance of the derivative's estimator corresponds to:

$$\mathrm{Var}\left[\left(\frac{\widehat{\mathrm{d}E\left(x\right)}}{\mathrm{d}x}\right)_{num}\right] = \frac{\mathrm{Var}\left[\widehat{E\left(x+\delta\right)}\right] + \mathrm{Var}\left[\widehat{E\left(x-\delta\right)}\right]}{4\delta^2}. \tag{8.19}$$

Moreover, when the energy estimator variances are assumed to be equal because δ is small, this equation can be rewritten as:

$$\mathrm{Var}\left[\left(\frac{\widehat{\mathrm{d}E\left(x\right)}}{\mathrm{d}x}\right)_{num}\right] = \frac{\mathrm{Var}\left[\widehat{E}\right]}{2\delta^2}, \tag{8.20}$$

where $\mathrm{Var}\left[\widehat{E}\right]$ is the variance of an energy estimator.

Since the estimator for the energy derivative is also a normally distributed random variable, ε_d was defined in the following manner:

$$\varepsilon_d = z_d \sqrt{\mathrm{Var}\left[\widehat{\left(\frac{\mathrm{d}E\left(x\right)}{\mathrm{d}x}\right)}_{num}\right]}, \tag{8.21}$$

where

$$z_d = \frac{\widehat{\left(\frac{\mathrm{d}E(x)}{\mathrm{d}x}\right)}_{num} - \left(\frac{\mathrm{d}E(x)}{\mathrm{d}x}\right)_{num}}{\sqrt{\mathrm{Var}\left[\widehat{\left(\frac{\mathrm{d}E(x)}{\mathrm{d}x}\right)}_{num}\right]}}. \tag{8.22}$$

Then, substituting the previous equation and eq. (8.14) into eq. (8.20) gives:

$$m = \left(\frac{\left(z_d\right)^2 M}{2\delta^2 \varepsilon_d^2}\right) \sum_{i=1}^{M} h_i^2 \mathrm{Var}\left[\langle \xi_t | \hat{P}_{z,i} | \xi_t \rangle\right]. \tag{8.23}$$

To complete the derivation, the total number of measurements m_d is defined as twice the number of measurements per energy value m and the inequality in eq. (8.15) is used. The outcome is the following inequality:

$$m_d \leq \left(\frac{\left(z_d\right)^2 M}{\delta^2 \varepsilon_d^2}\right) \sum_{i=1}^{M} h_i^2. \tag{8.24}$$

This result shows that it is possible to reduce the number of measurements required for a given uncertainty by increasing the step size. However, increasing the step size will also increase the error in the numerical derivative. This error corresponds to $O\left(\delta^2\right)$ [102].

References

[1] E. Schrödinger, Ann. Phys. (Leipzig) **384**, 361 (1926).

[2] D. R. Hartree, Math. Proc. Cambridge **24**, 89 (1928).

[3] P. A. M. Dirac, Proc. R. Soc. A **123**, 714 (1929).

[4] I. N. Levine, *Quantum Chemistry*, 7th ed. (Pearson Education, Upper Saddle River, NJ, 2014) p. 700.

[5] G. Fano and S. Blinder, *Twenty-First Century Quantum Mechanics: Hilbert Space to Quantum Computers - Mathematical Methods and Conceptual Foundations* (Springer International Publishing, Cham, Switzerland, 2017).

[6] D. Marx and H. Jürg, *Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods* (Cambridge University Press, Cambridge, 2009).

[7] C. C. J. Roothaan, Rev. Mod. Phys. **23**, 69 (1951).

[8] E. Schrödinger, Ann. Phys. (Leipzig) **385**, 437 (1926).

[9] C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

[10] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[11] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[12] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons, Chichester, England, 2000).

[13] A. J. Cohen, P. Mori-Sánchez, and W. Yang, Chem. Rev. **112**, 289 (2012).

[14] R. P. Feynman, Int. J. Theor. Phys. **21**, 467 (1982).

[15] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2010).

[16] D. Deutsch, Proc. R. Soc. A **400**, 97 (1985).

[17] D. P. DiVincenzo, Fortschr. Phys. **48**, 771 (2000).

[18] J. Preskill, Proc. R. Soc. A **454**, 385 (1998).

[19] A. G. Fowler, M. Mariantoni, J. M. Martinis, and A. N. Cleland, Phys. Rev. A **86**, 032324 (2012).

[20] M. Gong, M.-C. Chen, Y. Zheng, S. Wang, C. Zha, H. Deng, Z. Yan, H. Rong, Y. Wu, S. Li, F. Chen, Y. Zhao, F. Liang, J. Lin, Y. Xu, C. Guo, L. Sun, A. D. Castellano, H. Wang, C. Peng, C.-Y. Lu, X. Zhu, and J.-W. Pan, Phys. Rev. Lett. **122**, 110501 (2019).

[21] B. P. Lanyon, C. Hempel, D. Nigg, M. Muller, R. Gerritsma, F. Zahringer, P. Schindler, J. T. Barreiro, M. Rambach, G. Kirchmair, M. Hennrich, P. Zoller, R. Blatt, and C. F. Roos, Science **334**, 57 (2011).

[22] R. Barends, L. Lamata, J. Kelly, L. García-Álvarez, A. G. Fowler, A. Megrant, E. Jeffrey, T. C. White, D. Sank, J. Y. Mutus, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, I.-C. Hoi, C. Neill, P. J. J. O'Malley, C. Quintana, P. Roushan, A. Vainsencher, J. Wenner, E. Solano, and J. M. Martinis, Nat. Commun. **6**, 7654 (2015).

[23] K. Chenoweth, A. C. T. van Duin, and W. A. Goddard, J Phys. Chem. A **112**, 1040 (2008).

[24] J. G. Lee, *Computational Materials Science: An Introduction* (CRC Press, Boca Raton, FL, 2017).

[25] D. M. Duffy, Phil. Trans. R. Soc. A **368**, 3315 (2010).

[26] C. Bungaro, C. Noguera, P. Ballone, and W. Kress, Phys. Rev. Lett. **79**, 4433 (1997).

[27] R. Puthenkovilakam, E. A. Carter, and J. P. Chang, Phys. Rev. B **69**, 155329 (2004).

[28] J. M. Frost, K. T. Butler, and A. Walsh, APL Mater. **2**, 081506 (2014).

[29] S. Piana, K. Lindorff-Larsen, and D. E. Shaw, Biophys. J. **100**, L47 (2011).

[30] S. Erbas-Cakmak, D. A. Leigh, C. T. McTernan, and A. L. Nussbaumer, Chem. Rev. **115**, 10081 (2015).

[31] R. Eelkema, M. M. Pollard, J. Vicario, N. Katsonis, B. S. Ramon, C. W. M. Bastiaansen, D. J. Broer, and B. L. Feringa, Nature **440**, 163 (2006).

[32] R. S. Stoll and S. Hecht, Angew. Chem. Int. Edit. **49**, 5054 (2010).

[33] M. J. S. Dewar, E. G. Zoebisch, E. F. Healy, and J. J. P. Stewart, J. Am. Chem. Soc. **107**, 3902 (1985).

[34] P. D. Lyne, M. Hodoscek, and M. Karplus, J. Phys. Chem. A **103**, 3462 (1999).

[35] P. Ehrenfest, Z. Phys. **45**, 455 (1927).

[36] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[37] S. Lloyd, Science **273**, 1073 (1996).

[38] S. Wiesner, arXiv:9603028 [quant-ph] (1996).

[39] C. Zalka, Proc. R. Soc. Lond. A **454**, 313 (1998).

[40] D. Coppersmith, *IBM Research Report No. RC19642*, Tech. Rep. (1994).

[41] D. A. Lidar and H. Wang, Phys. Rev. E **59**, 2429 (1999).

[42] I. Kassal, S. P. Jordan, P. J. Love, M. Mohseni, and A. Aspuru-Guzik, Proceedings of the National Academy of Sciences of the United States of America 10.1073/pnas.0808245105 (2008), arXiv:0008033 [quant-ph] .

[43] D. S. Abrams and S. Lloyd, Phys. Rev. Lett. **83**, 5162 (1999).

[44] A. Aspuru-Guzik, A. D. Dutoi, P. J. Love, and M. Head-Gordon, Science **309**, 1704 (2005).

[45] J. R. McClean, R. Babbush, P. J. Love, and A. Aspuru-Guzik, J. Phys. Chem. Lett. **5**, 4368 (2014).

[46] M. B. Hastings, D. Wecker, B. Bauer, and M. Troyer, Quantum Inf. Comput. **15**, 1 (2015).

[47] Y. Li and S. C. Benjamin, Phys. Rev. X **7**, 21050 (2017).

[48] P. J. J. O'Malley, R. Babbush, I. D. Kivlichan, J. Romero, J. R. McClean, R. Barends, J. Kelly, P. Roushan, A. Tranter, N. Ding, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. G. Fowler, E. Jeffrey, E. Lucero, A. Megrant, J. Y. Mutus, M. Neeley, C. Neill, C. Quintana, D. Sank, A. Vainsencher, J. Wenner, T. C. White, P. V. Coveney, P. J. Love, H. Neven, A. Aspuru-Guzik, and J. M. Martinis, Phys. Rev. X **6**, 31007 (2016).

[49] E. Knill, G. Ortiz, and R. D. Somma, Phys. Rev. A **75**, 012328 (2007).

[50] K. M. Svore, M. B. Hastings, and M. Freedman, Quant. Inf. Comp. **14**, 306 (2013).

[51] B. P. Lanyon, J. D. Whitfield, G. G. Gillett, M. E. Goggin, M. P. Almeida, I. Kassal, J. D. Biamonte, M. Mohseni, B. J. Powell, M. Barbieri, A. Aspuru-Guzik, and A. G. White, Nat. Chem. **2**, 106 (2010).

[52] S. Paesani, A. A. Gentile, R. Santagati, J. Wang, N. Wiebe, D. P. Tew, J. L. O'Brien, and M. G. Thompson, Phys. Rev. Lett. **118**, 100503 (2017).

[53] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, Nat. Commun. **5**, 4213 (2014).

[54] J. R. McClean, J. Romero, R. Babbush, and A. Aspuru-Guzik, New J. Phys. **18**, 023023 (2016).

[55] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow, and J. M. Gambetta, Nature **549**, 242 (2017).

[56] Y. Shen, X. Zhang, S. Zhang, J.-N. Zhang, M.-H. Yung, and K. Kim, Phys. Rev. A **95**, 020501 (2017).

[57] C. Hempel, C. Maier, J. Romero, J. McClean, T. Monz, H. Shen, P. Jurcevic, B. P. Lanyon, P. Love, R. Babbush, A. Aspuru-Guzik, R. Blatt, and C. F. Roos, Phys. Rev. X **8**, 031022 (2018).

[58] Y. Nam, J.-S. Chen, N. C. Pisenti, K. Wright, C. Delaney, D. Maslov, K. R. Brown, S. Allen, J. M. Amini, J. Apisdorf, K. M. Beck, A. Blinov, V. Chaplin, M. Chmielewski, C. Collins, S. Debnath, A. M. Ducore, K. M. Hudek, M. Keesan, S. M. Kreikemeier, J. Mizrahi, P. Solomon, M. Williams, J. D. Wong-Campos, C. Monroe, and J. Kim, arXiv:1902.10171 (2019).

[59] M. Born and K. Huang, *Oxford University Press* (1962).

[60] J. Goodisman, *Diatomic Interaction Potential Theory: Volume 1 - Fundamentals* (Academic Press, New York, 1973).

[61] W. Heisenberg, Z. Phys. **43**, 172 (1927).

[62] J. J. Sakurai, *Modern Quantum Mechanics: Revised Edition* (Addison-Wesley Publishing Company, Reading, Massachusetts, 1994).

[63] P. J. Mohr, D. B. Newell, and B. N. Taylor, Rev. Mod. Phys. 10.1103/RevModPhys.88.035009 (2016).

[64] M. Born and R. Oppenheimer, Ann. Phys. (Leipzig) **389**, 457 (1927).

[65] G. L. Closs and J. R. Miller, Science **240**, 440 (1988).

[66] N. Bohr, Z. Phys. **2**, 423 (1920).

[67] D. Marx, Chem. Phys. Chem. **7**, 1848 (2006).

[68] N. L. Doltsinis and D. Marx, J.Theor. Comput. Chem. **01**, 319 (2002).

[69] D. R. Hartree, W. Hartree, and B. Swirles, Phil. Trans. R. Soc. A **238**, 229 (1939).

[70] B. O. Roos, P. R. Taylor, and P. E. Siegbahn, Chem. Phys. **48**, 157 (1980).

[71] W. Pauli, Z. Phys. **31**, 765 (1925).

[72] J. C. Slater, Phys. Rev. **34**, 1293 (1929).

[73] M. L. Abrams and C. D. Sherrill, J. Chem. Phys. **118**, 1604 (2003).

[74] J. A. Pople and R. K. Nesbet, J. Chem. Phys. **22**, 571 (1954).

[75] L. E. McMurchie and E. R. Davidson, J. Comp. Phys. **26**, 218 (1978).

[76] W. J. Hehre, R. F. Stewart, and J. A. Pople, J. Chem. Phys. **51**, 2657 (1969).

[77] F. Jensen, *Introduction to Computational Chemistry* (John Wiley & Sons, Chichester, UK, 2017).

[78] A. Szabo and N. S. Ostlund, *McGraw-Hili, New York* (Dover Publications, Mineola, NY, 1996).

[79] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

[80] P. Pulay, J. Comp. Chem. **3**, 556 (1982).

[81] P. R. Surjan, *Second Quantized Approach to Quantum Chemistry: An Elementary Introduction* (Springer-Verlag, Berlin, 1989).

[82] K. Nakata, T. Murase, T. Sakuma, and T. Takada, J. Comput. Appl. Math. **149**, 351 (2002).

[83] W. Kutzelnigg, in *Methods of Electronic Structure Theory*, edited by H. F. Schaefer (Springer Science+Business Media, New York, 1977) pp. 129–188.

[84] J. Romero, R. Babbush, J. R. McClean, C. Hempel, P. J. Love, and A. Aspuru-Guzik, Quantum Sci. Technol. **4**, 014008 (2018).

[85] M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).

[86] P. Jordan and E. Wigner, Z. Phys. **47**, 631 (1928).

[87] R. Somma, G. Ortiz, J. E. Gubernatis, E. Knill, and R. Laflamme, Phys. Rev. A **65**, 042323 (2002).

[88] J. T. Seeley, M. J. Richard, and P. J. Love, J. Chem. Phys. **137**, 224109 (2012).

[89] A. Tranter, S. Sofia, J. Seeley, M. Kaicher, J. McClean, R. Babbush, P. V. Coveney, F. Mintert, F. Wilhelm, and P. J. Love, Int. J. Quantum Chem. **115**, 1431 (2015).

[90] B. C. Hall, in *Lie Groups, Lie Algebras, and Representations* (Springer International Publishing, Cham, 2015) pp. 109–137.

[91] G. Aleksandrowicz, T. Alexander, P. Barkoutsos, L. Bello, Y. Ben-Haim, D. Bucher, F. J. Cabrera-Hernández, J. Carballo-Franquis, A. Chen, C.-F. Chen, J. M. Chow, A. D. Córcoles-Gonzales, A. J. Cross, A. Cross, J. Cruz-Benito, C. Culver, S. D. L. P. González, E. D. L. Torre, D. Ding, E. Dumitrescu, I. Duran, P. Eendebak, M. Everitt, I. F. Sertage, A. Frisch, A. Fuhrer, J. Gambetta, B. G. Gago, J. Gomez-Mosquera, D. Greenberg, I. Hamamura, V. Havlicek, J. Hellmers, Ł. Herok, H. Horii, S. Hu, T. Imamichi, T. Itoko, A. Javadi-Abhari, N. Kanazawa, A. Karazeev, K. Krsulich, P. Liu, Y. Luh, Y. Maeng, M. Marques, F. J. Mart\'\in-Fernández, D. T. McClure, D. McKay, S. Meesala, A. Mezzacapo, N. Moll, D. M. Rodr\'\iguez, G. Nannicini, P. Nation, P. Ollitrault, L. J. O'Riordan, H. Paik, J. Pérez, A. Phan, M. Pistoia,

V. Prutyanov, M. Reuter, J. Rice, A. R. Davila, R. H. P. Rudy, M. Ryu, N. Sathaye, C. Schnabel, E. Schoute, K. Setia, Y. Shi, A. Silva, Y. Siraichi, S. Sivarajah, J. A. Smolin, M. Soeken, H. Takahashi, I. Tavernelli, C. Taylor, P. Taylour, K. Trabing, M. Treinish, W. Turner, D. Vogt-Lee, C. Vuillot, J. A. Wildstrom, J. Wilson, E. Winston, C. Wood, S. Wood, S. Wörner, I. Y. Akhalwaya, and C. Zoufal, Qiskit: An Open-source Framework for Quantum Computing (2019).

[92] J. D. Whitfield, J. Biamonte, and A. Aspuru-Guzik, Mol. Phys. **109**, 735 (2011).

[93] D. Wecker, M. B. Hastings, and M. Troyer, Phys. Rev. A **92**, 042303 (2015).

[94] J. Spall, IEEE T. Automat. Contr. **37**, 332 (1992).

[95] J. Spall, IEEE T. Aero. Elec. Sys. **34**, 817 (1998).

[96] J. Poulson, B. Marker, R. A. van de Geijn, J. R. Hammond, and N. A. Romero, ACM T. Math. Software **39**, 1 (2013).

[97] K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules* (Springer Science+Business Media, New York, 1979).

[98] M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, G. A. Petersson, H. Nakatsuji, X. Li, M. Caricato, A. V. Marenich, J. Bloino, B. G. Janesko, R. Gomperts, B. Mennucci, H. P. Hratchian, J. V. Ortiz, A. F. Izmaylov, J. L. Sonnenberg, D. Williams-Young, F. Ding, F. Lipparini, F. Egidi, J. Goings, B. Peng, A. Petrone, T. Henderson, D. Ranasinghe, V. G. Zakrzewski, J. Gao, N. Rega, G. Zheng, W. Liang, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, K. Throssell, J. A. Montgomery Jr., J. E. Peralta, F. Ogliaro, M. J. Bearpark, J. J. Heyd, E. N. Brothers, K. N. Kudin, V. N. Staroverov, T. A. Keith,

R. Kobayashi, J. Normand, K. Raghavachari, A. P. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, J. M. Millam, M. Klene, C. Adamo, R. Cammi, J. W. Ochterski, R. L. Martin, K. Morokuma, O. Farkas, J. B. Foresman, and D. J. Fox, Gaussian 16 Revision B.01 (2016).

[99] A. K. Wilson, T. van Mourik, and T. H. Dunning, J. Mol. Struc.: THEOCHEM **388**, 339 (1996).

[100] R. A. Kendall, T. H. Dunning, and R. J. Harrison, J. Chem. Phys. **96**, 6796 (1992).

[101] J. A. Pople, Reviews of Modern Physics **71**, 1267 (1999).

[102] B. Fornberg, Math. Comput. **51**, 699 (1988).

[103] W. C. Swope, H. C. Andersen, P. H. Berens, and K. R. Wilson, J. Chem. Phys. **76**, 637 (1982).

[104] D. Frenkel and B. Smit, *Understanding Molecular Simulation: From Algorithms to Applications* (Academic Press, San Diego, 2002).

[105] V. Wathelet, B. Champagne, D. H. Mosley, J.-M. André, and S. Massidda, Chemical Physics Letters **275**, 506 (1997).

[106] G. Strang, *Linear Algebra and its Applications* (Thomson Brooks/Cole, Belmont, CA, 2006).

[107] M. Sipser, *Introduction to the Theory of Computation* (Thomson Course Technology, Boston, 2006).

[108] D. DiVincenzo and B. Terhal, Phys. World **11**, 53 (1998).

[109] S. B. Bravyi and A. Y. Kitaev, Ann. Phys. **298**, 210 (2002).

[110] T. Brecht, W. Pfaff, C. Wang, Y. Chu, L. Frunzio, M. H. Devoret, and R. J. Schoelkopf, npj Quantum Information **2**, 16002 (2016).

[111] R. Barends, J. Kelly, A. Megrant, A. Veitia, D. Sank, E. Jeffrey, T. C. White, J. Mutus, A. G. Fowler, B. Campbell, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, C. Neill, P. O'Malley, P. Roushan, A. Vainsencher, J. Wenner, A. N. Korotkov, A. N. Cleland, and J. M. Martinis, Nature **508**, 500 (2014).

[112] K. Temme, S. Bravyi, and J. M. Gambetta, Physical Review Letters **119**, 10.1103/PhysRevLett.119.180509 (2017), arXiv:1612.02058 .

[113] T. E. O'Brien, B. Senjean, R. Sagastizabal, X. Bonet-Monroig, A. Dutkiewicz, F. Buda, L. DiCarlo, and L. Visscher, arXiv:1905.03742 (2019).

[114] K. Mitarai, Y. O. Nakagawa, and W. Mizukami, arXiv:1905.04054 (2019).

[115] T. Tamayo-Mendoza, C. Kreisbeck, R. Lindh, and A. Aspuru-Guzik, ACS Central Science **4**, 559 (2018).

[116] M. H. DeGroot and M. J. Schervish, *Probability and Statistics* (Pearson Education, Boston, 2012).

[117] S. L. Miller and D. Childers, *Probability and Random Processes: With Applications to Signal Processing and Communications: Second Edition* (Elsevier, Waltham, MA, 2012).

[118] D. S. Lemons and P. Langevin, *An Introduction to Stochastic Processes in Physics : Containing on the Theory of Brownian Motion by Paul Langevin, Translated by Anthony Gythiel* (John Hopkins University Press, Baltimore, 2002).