BACHELOR

Quantum chemistry and the Variational Quantum Eigensolver

Bogaarts, Timo J.

Award date:
2019

Disclaimer
This document contains a student thesis (bachelor's or master's), as authored by a student at Eindhoven University of Technology. Student theses are made available in the TU/e repository upon obtaining the required degree. The grade received is not published on the document as presented in the repository. The required complexity or quality of research of student theses may vary by program, and the required minimum study period may vary in duration.

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
Quantum Chemistry and the Variational Quantum Eigensolver

CQT 2019-12
Abstract
Quantum Computation promises very large speed ups with respect to normal, classical computers on specific problems. Examples include prime-factoring and database searching. These applications are far away however, since they require high-precision logic gates, which are not realisable on today’s quantum devices. These quantum devices can be used for other purposes however, one of which is a Variational Quantum Eigensolver (VQE). This VQE will, using a variational method, try to obtain the lowest eigenvalue of a specific Hamiltonian.

In this thesis the VQE was simulated on a classical computer and applied to the quantum chemistry problem of calculating the ground state of small molecules (H₂, LiH, F₂). A brief introduction to both quantum chemistry and quantum computing is also given. Two different trial state methods are investigated: one which takes advantage of the hardware to efficiently prepare a trial state and one which modifies a standard theoretical quantum chemistry technique for use on a quantum computer. A detailed protocol is given for the implementation of this theoretical trial state on any quantum computer. Both trial state methods yield good results for the smallest of molecules but for somewhat larger molecules the theoretical trial state method shows better results. This method comes at a much larger computational cost however, and the specific scaling is investigated.

Methods to improve the accuracy and speed of the VQE with the hardware efficient trial state are investigated. These methods lead to a better estimation of the ground state energy but for larger molecules the improvement was minimal. However, various avenues of further improvement can readily be given, which include modifying the optimization algorithm and increasing the size of the quantum circuit.
## Contents

1 Introduction 4

2 Quantum Chemistry and classical methods 5
   2.1 The quantum chemistry problem 5
   2.2 The variational principle 5
   2.3 Many-electron wave functions 6
   2.4 Spin and positional coordinates 6
   2.5 Molecular spin-orbitals and basis sets 7
   2.6 Evaluating the Hamiltonian and second quantization 9
   2.7 The Hartree-Fock method 10
      2.7.1 The Hartree-Fock equations 11
      2.7.2 The Roothaan Equations 12
   2.8 Configuration Interaction 13
   2.9 Coupled Cluster methods 14

3 Quantum Computing 16
   3.1 Basics of Quantum Computation 16
   3.2 Entanglement and superposition 18
   3.3 Hamiltonian simulation 19
   3.4 Mapping Creation and Annihilation operators to qubit operations 21
      3.4.1 Jordan-Wigner transform 21
      3.4.2 Bravyi-Kitaev transform 23
   3.5 Qubit Reduction using locked spin-orbitals 23

4 The VQE algorithm 25
   4.1 Preparing a quantum state 25
      4.1.1 Hardware Tailored Ansatz 26
      4.1.2 Unitary Coupled Cluster 26
   4.2 Measuring the Hamiltonian 27
   4.3 Optimizing the variational parameters 28

5 Methods 30
   5.1 General Methods 30
   5.2 Hardware Tailored Ansatz methods 30
   5.3 Unitary Coupled Cluster methods 32
      5.3.1 Exact implementation 32
      5.3.2 Realistic Implementation 33
   5.4 Potential Curve methods 37

6 Results 38
   6.1 Hardware Tailored Ansatz Results 38
      6.1.1 The Hydrogen molecule 38
      6.1.2 The LiH molecule 40
1 Introduction

In this age computers form the basis of our society. They are used for a very wide range of applications and have become very fast due to miniaturization of the transistor. Their application also includes physics research, where some problems can be efficiently solved on them. Examples are computational General Relativity and computational fluid dynamics. However, some problems quickly become intractable because of the exponential scaling of the computational cost with system size. Factoring large numbers into primes is an example of a problem which scales exponentially with the size of the problem. Some problems which have such scaling are much more efficiently solved on a different type of computer: a quantum computer.

Originally proposed by Feynman in 1982[1] for the purpose of simulating quantum physics more efficiently these devices can actually be used for much wider purpose. Using an algorithm invented by Shor[2] quantum computers can factorize large numbers into primes in only polynomial time. However, a quantum computer that can solve this problem faster than a classical computer (quantum supremacy) has not yet been demonstrated due to strict requirements on the quantum computer. An algorithm that takes advantage of both quantum computers and classical computers was proposed [3] to lower these requirements and thus sooner prove quantum supremacy for a problem.

This algorithm is the Variational Quantum Eigensolver. It seeks to find the lowest eigenvalue of a Hamiltonian using a variational method. The quantum computer is used to represent the trial state, which would be inefficient on a classical computer [4]. The classical computer then runs an optimization algorithm which should converge to the lowest eigenvalue. This VQE has been applied to the calculation of the ground state of several small molecules, using a variety of trial states and optimization methods [3][5][6][7].

In this thesis a short introduction to quantum chemistry is given in section 2 which treats the topics relevant for the rest of the thesis. Also, a very brief introduction to quantum computation is given in section 3. The VQE algorithm is then outlined and the several steps in the algorithm are examined in detail in section 4. The implementation of two trial states in the VQE is discussed in section 5 while the results are discussed in section 6. A conclusion is given in section 7 and finally a outlook for further research is given in section 8.
2 Quantum Chemistry and classical methods

The following sections briefly introduce the important concepts of quantum chemistry needed for understanding the application of the VQE to it. The information in these sections is from the book "Modern Quantum Chemistry" by Szabo [8] unless stated otherwise.

2.1 The quantum chemistry problem

In quantum chemistry one is usually interested in the properties of ground states of molecules. These ground states obey the Schrödinger Equation

\[ \hat{H} \psi = E \psi, \]  

(2.1)

with \( \hat{H} \) defined in the many-electron and many-nuclei case as [9]

\[ \hat{H} = - \sum_i \frac{\nabla^2 \bar{R}_i}{2M_i} - \sum_i \frac{\nabla^2 \bar{r}_i}{2} - \sum_{i,j} \frac{Z_i}{|\bar{R}_i - \bar{r}_j|} + \frac{1}{2} \sum_{i,j} \frac{Z_i Z_j}{|\bar{R}_i - \bar{R}_j|} + \frac{1}{2} \sum_{i,j} \frac{1}{|\bar{r}_i - \bar{r}_j|}, \]  

(2.2)

In the above equation \( \bar{R}_i \) are the positions of the nuclei, \( Z_i \) are the charges of the nuclei \( M_i \) are the masses of the nuclei and \( \bar{r}_i \) are the positions of the electrons. Also, atomic units are used, i.e. \( \frac{1}{4\pi \varepsilon_0} = m_e = q_e = \hbar = 1 \). The first term concerns the kinetic energy of the nuclei, the second the kinetic energy of the electrons, the third electron-nuclei potential energy, the fourth the nuclei-nuclei potential energy and the fifth the electron-electron potential energy.

Since the mass of the electron is much smaller than the mass of the nuclei (even for the hydrogen atom the ratio is \( m_p/m_e = 1835 \) the dynamics of the nuclei will be neglected (the nuclei will thus remain at fixed positions), which is the Oppenheimer-Born approximation. The first term will then be equal to zero and the Hamiltonian simplifies to the more tractable

\[ \hat{H} = - \sum_i \frac{\nabla^2 \bar{r}_i}{2} + \frac{1}{2} \sum_{i,j} \frac{1}{|\bar{r}_i - \bar{r}_j|} - \sum_{i,j} \frac{Z_i}{|\bar{R}_i - \bar{r}_j|} + V_{\text{nuclei}}, \]  

(2.3)

where \( V_{\text{nuclei}} \) was introduced as shorthand for the fourth term in equation (2.2) and is independent of the electron configuration. The Hamiltonian can then be subdivided in the one-electron operators of the first and third term, the two-electron operator of the second term and a constant.

2.2 The variational principle

The eigenvalue equation (2.1) cannot be solved exactly. To obtain an approximate value for the lowest eigenvalue (in other words, the ground state energy of the Hamiltonian) the variational principle can be used. This principle states that for any normalized state \( \psi \) it holds that

\[ E_g \leq \langle \psi | \hat{H} | \psi \rangle. \]  

(2.4)
If a state is parameterized by certain parameters the ground state energy may thus be approximated by minimizing the expectation value of the Hamiltonian on that state under the condition that it remains normalized.

2.3 Many-electron wave functions

Since electrons have half-integer spins they are fermions and thus obey the antisymmetry principle. This principle states that electrons are indistinguishable, and when these two particles are interchanged the wave function should change sign. In other words, if $\Psi$ is $n$-many-electron wave function depending on $n$ positional coordinates, it should hold that

$$|\Psi(\vec{r}_1, \vec{r}_2, ..., \vec{r}_n)\rangle = -|\Psi(\vec{r}_2, \vec{r}_1, ..., \vec{r}_n)\rangle.$$  \hspace{1cm} (2.5)

In quantum chemistry it is customary to work with single-electron wave functions. These functions can for example be a Linear Combination of Atomic Orbitals (LCAO). Antisymmetry can then be ensured by using a determinant, called a Slater determinant. In a system with $N$ electrons described by the $N$ coordinates $\{x_i\}$ and $N$ single-electron wave functions $\{\phi_j\}$, the total wave function can thus be described as

$$|\Psi(x_1, x_1, ..., x_N)\rangle = \frac{1}{N!} \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & ... & \phi_N(x_1) \\ \phi_1(x_2) & \phi_2(x_2) & ... & \phi_N(x_2) \\ ... & ... & ... & ... \\ \phi_1(x_N) & \phi_2(x_N) & ... & \phi_N(x_N) \end{vmatrix}.$$  \hspace{1cm} (2.6)

The interchange of two positional coordinates is in this form equivalent to exchanging columns of the determinant. The determinant changes sign under column exchange and thus the antisymmetric principle is obeyed. Written in this form it also becomes clear that two electrons cannot be in the same single-electron wave function since the determinant of a matrix with two identical columns is zero. The antisymmetry principle thus also implies the Pauli exclusion principle.

The determinant is completely specified up to a sign by the collection of single-electron wave functions. It is then convenient to notate such a determinant as

$$|\Psi\rangle = |\phi_1 \phi_2 ... \phi_n\rangle.$$  \hspace{1cm} (2.7)

2.4 Spin and positional coordinates

The state of the electron has to be specified by not only position, but also spin. The total single-electron wave function then becomes

$$\phi(x) = \phi(r, \omega) = \psi(r)\alpha(\omega) \text{ or } \psi(r)\beta(\omega)$$  \hspace{1cm} (2.8)

where $x$ denotes a combined positional and spin coordinate, $\psi(r)$ denotes a spatial wave function depending on the positional coordinate $r$, and $\{\alpha(\omega), \beta(\omega)\}$ denote two orthonormal spin functions depending on a spin coordinate $\omega$. Orthonormality implies that

$$\int d\omega \, \alpha(\omega)\alpha(\omega) = \int d\omega \, \beta(\omega)\beta(\omega) = 1 \text{ and } \int d\omega \, \alpha(\omega)\beta(\omega) = 0.$$  \hspace{1cm} (2.9)

Adding the spin onto the spatial wave function now allows two electrons having different spins to reside in the same spatial wave function.
2.5 Molecular spin-orbitals and basis sets

In quantum chemistry the Linear Combination of Atomic Orbitals (LCAO) technique is often employed. In LCAO \( K \) atomic spatial orbitals (AO’s), such as the 1s hydrogen orbital, are combined to form \( K \) spatial Molecular Orbitals (MO’s). For example, in the hydrogen molecule the 1s orbitals of both atoms are combined to form two molecular orbitals as

\[
\psi_{MO\pm} = \alpha \psi_{1s,1} \pm \beta \psi_{1s,2}
\]

where \( \psi_{MO\pm} \) are the two molecular spatial orbitals, \( \alpha \) and \( \beta \) are two constants indicating the weight of the atomic orbital in the molecular orbital, and \( \{\psi_{1s,1}, \psi_{1s,2}\} \) are the 1s spatial orbitals of hydrogen atom 1 and 2. This process is shown in figure [2.1].

![Figure 2.1: The process of creating molecular orbitals from atomic orbitals for molecular hydrogen. The height difference corresponds to energy difference, calculated using the Hartree-Fock method. Image adapted from [10].](image)

More general,

\[
\psi_{MO,j} = \sum_i c_{ij} \psi_{AO,i}.
\]  

The spin function has to be added onto these molecular orbitals to obtain molecular spin-orbitals, denoted as \( \{X_j\} \). Note that in general there are more molecular spin-orbitals \( K \) than electrons \( N \) in the system and a subset (or combinations of subsets) of those spin-orbitals should be selected for the groundstate.

The atomic orbitals that are included depend on the basis set used, of which there are several. A minimal basis set compromises only the orbitals that are occupied in the neutral atom, so for hydrogen only the 1s orbital is included, while for carbon the 1s, 2s, 2p<sub>x</sub>, 2p<sub>y</sub>, and 2p<sub>z</sub> orbitals are included.\(^\text{1}\) Using more atomic orbitals leads to higher accuracies since a larger part of the Hilbert space is taken into account, but comes at computational costs. Double-zeta basis sets use two atomic orbitals for every occupied orbital, and various other basis sets have been published, such as cc-pVDZ. Another useful trick is to include more

\(^1\)Note that in this case there is an extra orbital which is not occupied but is still included due to symmetry considerations.
atomic orbitals for the valence electrons than for the core electrons, which is also what the cc-pVDZ basis set does.

The Hartree-Fock method, which will be described in detail in later sections, then calculates, using an approximation, the optimal coefficients $c_{ij}$ such that

$$\langle \chi_1 \chi_2 ... \chi_N | \hat{H} | \chi_1 \chi_2 ... \chi_N \rangle$$

is minimized. The notation of equation (2.7) was used for the Slater determinant and $N$ is the number of electrons. In the Hartree-Fock method the groundstate only involves the $N$ lowest energy orbitals while it calculates all $K$ molecular orbitals. A combination of different sets of spin-orbitals is in principle also a possibility. Such combinations are the topic of several post-Hartree-Fock methods which seek to improve the approximation of the ground state energy and some of them will be discussed in later sections.

To do actual calculates on these orbitals they have to be represented in a explicit functional form. The main requirements for these functions are: needing only a small amount of atomic orbitals for an accurate representation of a molecular orbitals and having a small computational time cost for integral evaluation. Slater type orbitals have radial functions of the form

$$R(r) = c \ r^{n-1} e^{-\zeta r},$$

where $c$ is a normalization constant, $n$ is the principal quantum number, $r$ is the distance from the nucleus, and $\zeta$ is a constant depending on the nucleus. These kind of orbitals satisfy the first requirement but fail at the second. Another method uses some amount of Gaussian functions to approximate a single Slater function, since Gaussian functions are easily integrable. An example of such a basis set is the STO-3G, which is a minimal basis set in which Slater type orbitals are approximated as

$$\psi_j = \sum_{i=1}^{3} k_{ij} G_i(\alpha_{ij}, r - R_n),$$

where $\psi_j$ is the $j$th atomic orbital, $k_{ij}$ and $\alpha_{ij}$ are constants that should be obtained by a least-squares fit to maximize the overlap between the three-Gaussian approximation and the actual Slater orbital, and $G_i$ is a normalized Gaussian function. The coefficients in this equation are thus not variational parameters and should be determined prior to beginning a Hartree-Fock calculation. To approximate different kinds of orbitals different kinds of Gaussian functions should be used. Three examples of normalized Gaussian functions for approximating the $1s$, $2p_x$ and $3d_{xy}$ Slater orbitals are

$$G_{1s}(\alpha, r) = \left( \frac{8\alpha^3}{\pi^3} \right)^{\frac{1}{4}} e^{-\alpha r^2},$$

$$G_{2p_x}(\alpha, r) = \left( \frac{128\alpha^5}{\pi^3} \right)^{\frac{1}{4}} x e^{-\alpha r^2},$$

$$G_{3d_{xy}}(\alpha, r) = \left( \frac{2048\alpha^7}{\pi^3} \right)^{\frac{1}{4}} xy e^{-\alpha r^2}.$$
2.6 Evaluating the Hamiltonian and second quantization

The energy of a system in a state $|\Psi\rangle$ is equal to $\langle\Psi|\hat{H}|\Psi\rangle$. It is also needed to evaluate some matrix elements like $\langle\Psi_1|\hat{H}|\Psi_2\rangle$. Separating the Hamiltonian of equation (2.3) into a constant, two single-electron operators and a double electron operator as

$$\hat{H}_0 = V_{\text{nuclei}},$$

$$\hat{H}_1 = -\sum_i \left( \frac{\nabla^2}{2} - \sum_j Z_i \frac{1}{|\vec{R}_i - \vec{r}_j|} \right),$$

$$\hat{H}_2 = \sum_{i>j} \frac{1}{|\vec{r}_i - \vec{r}_j|},$$

and introducing a shorthand notation for the Slater determinant by only notating the indices of the spin-orbitals as

$$|\Psi_1\rangle = |a_b...\rangle = |ab...\rangle$$

$$|\Psi_2\rangle = |cd...\rangle,$$

these matrix elements can be concisely written as

$$\langle\Psi_1|\hat{H}|\Psi_1\rangle = V_{\text{nuclei}} + \langle ab...|\hat{H}_1|ab...\rangle + \langle ab...|\hat{H}_2|ab...\rangle.$$

The $\langle\alpha\beta|\hat{H}_i|\gamma\delta\rangle$ terms involve integrals over the Slater determinants and can become very complex. For example, the $\langle12|\hat{H}_1|12\rangle$ element for the minimal basis set for the hydrogen molecule (only two 1s orbitals) is

$$\langle12|\hat{H}_1|12\rangle = 2 \times \int dx_1 dx_2 \frac{1}{2} [\chi_1(x_1)\chi_2(x_2) - \chi_2(x_1)\chi_1(x_2)]^* \times \hat{H}_1 [\chi_1(x_1)\chi_2(x_2) - \chi_2(x_1)\chi_1(x_2)]$$

$$= \int dx_1 \chi_1(x_1)^* \hat{H}_1 \chi_1(x_1) + \int dx_1 \chi_2(x_1)^* \hat{H}_1 \chi_2(x_1)$$

$$= \langle1|\hat{H}_1|1\rangle + \langle2|\hat{H}_1|2\rangle = h_{11} + h_{22}.$$

The factor 2 arises from the summation over $i$ in equation (2.18), the factor $\frac{1}{2}$ from the normalization of the determinants, the second step follows from an explicit calculation taking into account the orthonormality of the spin-orbitals, and the last step introduces yet another new notation. This notation is also used in a more general sense for both one- and two-electron operators as

$$h_{pq} = \int dx \chi_p^*(x)(\frac{\nabla^2}{2} - \sum_i \frac{Z_i}{|\vec{R}_i - \vec{r}|})\chi_q(x),$$

$$h_{pqrs} = \int dx_1 dx_2 \frac{\chi_p^*(x_1)\chi_q^*(x_2)\chi_s(x_1)\chi_r(x_2)}{|r_1 - r_2|}. $$

9
The two-electron integral has two different types which each have their distinct physical interpretation. The $h_{abba}$ type are just the coulomb repulsion between two charge clouds

$$h_{abba} = \int d\mathbf{x}_1 \, d\mathbf{x}_2 \frac{|\chi_a(\mathbf{x}_1)|^2 |\chi_b(\mathbf{x}_2)|^2}{\vec{r}_{12}}, \quad (2.24)$$

while the $h_{abab}$ type results from the antisymmetry principle and are called the exchange integrals. There exists rules on how to calculate matrix elements in terms of the above integrals instead of carrying out an explicit calculation with large determinants, but these are beyond the scope of this discussion.

There exists another formalism which incorporates those integrals more centrally, called the second quantization method. This formalism introduces creation and annihilation operators $\hat{a}^\dagger$ and $\hat{a}$. The creation operator adds a spin-orbital to the Slater determinant, while the annihilation operator removes a spin-orbital. If a fermionic vacuum state (which contains no electrons and thus no occupied spin-orbitals) is defined as $|\rangle$ these operators obey the following properties

$$\begin{align*}
(\hat{a}^\dagger_j)^\dagger &= \hat{a}_j, \quad \hat{a}^\dagger_j |j\rangle &= |X_j\rangle = |j\rangle, \quad \hat{a}_j |j\rangle &= 0, \quad \hat{a}^\dagger_j |kln\rangle &= |j kln\rangle, \\
\hat{a}_j |kln\rangle &= |kln\rangle, \quad \hat{a}_j |kln\rangle &= -\hat{a}_j |kln\rangle = -|kln\rangle, \\
\hat{a}^\dagger_j |kln\rangle &= 0, \quad \hat{a}_j |kln\rangle &= 0, \quad \hat{a}_j \hat{a}_i + \hat{a}_i \hat{a}_j = \{\hat{a}_j, \hat{a}_i\} = 0 = \{\hat{a}^\dagger_j, \hat{a}_i^\dagger\} \quad (2.25) \\
\{\hat{a}_i, \hat{a}^\dagger_j\} &= \delta_{ij}. \quad (2.26)
\end{align*}$$

These relations convey that the adjoint of the creation operator is the annihilation operator and that the creation operator adds a spin-orbital to the left-most position on the Slater determinant. The annihilation operator can only remove the spin-orbital in the left-most position and columns in the determinant should be exchanged if the operator’s index does not correspond to the left-most index. The creation operator applied to a determinant that already has its spin-orbital is necessarily zero since the determinant then has two equal columns. The annihilation operator applied to a determinant which does not have its corresponding spin-orbital is also zero. The last few relations are the anti-commutators of the operators and are referred as the fermionic anticommutation relations. The Hamiltonian then takes the quite simple form

$$\hat{H} = \sum_{pq} h_{pq} \hat{a}^\dagger_p \hat{a}_q + \sum_{pqr} h_{pqr} \hat{a}^\dagger_p \hat{a}^\dagger_q \hat{a}_r \hat{a}_s \quad (2.29)$$

and it is this form that is mostly used in quantum chemistry computing.

### 2.7 The Hartree-Fock method

The goal of the Hartree-Fock (HF) method is to obtain expressions for the $K$ molecular orbitals. The HF-method assumes that the wave-function is only a single Slater determinant. This assumption implies that if one electron is measured to be in some orbital of this determinant, every other orbital in that determinant will still be always occupied by the other electrons. In reality, if an orbital is measured to be occupied other orbitals that are,
on average, close to this occupied orbital will have a lower chance to be occupied due to the coulomb repulsion. This single-determinant approximation thus neglects electron correlation. The Hamiltonian is not approximated, it is only the electron correlation that is neglected in the HF-method. This method provides the starting point for almost all other methods in quantum chemistry, which are collectively referred to as post-Hartree-Fock methods.

2.7.1 The Hartree-Fock equations

The Hartree-Fock method seeks to find a set of \( N \) spin-orbitals \( \{ \chi_i \} \) which minimizes \( E_0 = \langle \chi_1 \chi_2 \ldots \chi_N | \hat{H} | \chi_1 \chi_2 \ldots \chi_N \rangle \) under the constraint that every spin-orbital is normalized and that they form an orthonormal set, i.e.

\[
\langle \chi_i | \chi_j \rangle = \delta_{ij}.
\]  

(2.30)

Using a functional variation method it can then be shown that this set of spin-orbitals satisfy a non-linear eigenvalue equation

\[
\hat{f} | \chi_i \rangle = E_i | \chi_i \rangle = \left[ \hat{h}_i + \sum_{j \neq i}^N \hat{J}_j - \hat{K}_j \right] | \chi_i \rangle,
\]  

(2.31)

where \( \hat{f} \) is the Fock-operator, \( E_i \) is the energy of the spin orbital \( \chi_i \), \( \hat{h} \) is the one-electron operator, and \( \hat{J}_j \) and \( \hat{K}_j \) are the coulomb and exchange operators. The summation runs only over the \( N \) lowest energy spin-orbitals, so if these are found generating extra unoccupied spin-orbitals is comparatively easy. The one-electron operator is defined as

\[
\hat{h}_i = -\frac{1}{2} \nabla_i^2 - \sum_A \frac{Z_A}{|r_i - R_A|},
\]  

(2.32)

where \( \nabla_i^2 \) is the Laplacian acting on the positional coordinates \( \vec{r}_i \) of spin-orbital \( \chi_i \) and the sum runs over all the nuclei in the system, with charges \( Z_A \) and coordinates \( \vec{R}_A \). Taking the expectation value of this operator is the same as calculating the integral in equation (2.22). The coulomb operator is defined as

\[
\hat{J}_j \chi_i(x) = \int dx_j \frac{|\chi_j|^2}{|r_i - r_j|} \chi_j(x_1).
\]  

(2.33)

where the integration runs over the combined positional and spin coordinate of spin-orbital \( \chi_j \). Physically, the operator represents the coulombic potential at the coordinate \( r_i \) arising from an occupied spin-orbital \( \chi_j \). If an expectation value of it is taken as \( \langle \chi_j \chi_i | \hat{J}_j | \chi_j \chi_i \rangle \) it is equivalent to to the \( h_{ijji} \) types of integrals defined in equation (2.24). The exchange operator is defined by its action on a spin-orbital \( \chi_i \) as

\[
\hat{K}_j \chi_i(x_1) = \int dx_2 \frac{\chi_j^*(x_2)\chi_i(x_2)}{|r_i - r_j|} \chi_j(x_1).
\]  

(2.34)
The difference between equations (2.33) and (2.34) is simply that the exchange operator exchanges the spin-orbitals \( \chi_j \) and \( \chi_i \) relative to the coulomb operator. This exchange makes the operator non-local however, since acting on the spin-orbital \( \chi_i \) depends on the value of the spin-orbital everywhere, instead of just at the position \( x_i \). The exchange operator has no clear physical interpretation but its expectation value is the \( h_{ijij} \) type of integral discussed in section 2.6.

The eigenvalue equation (2.31) is non-linear, non-local and depends on not only the spin-orbital considered but on \( N-1 \) other spin-orbitals. Solving these equations either symbolically or numerically is very hard. This motivates the introduction of a basis set as defined in section 2.5 to transform these equations into matrix equations which are generally quite easy to solve. If this basis would be complete, the solution of these matrix equations would be equal to the exact solution of equation (2.31) and the energies gained from this are the Hartree-Fock limit. In practice a complete basis is never used and a finite amount of atomic spin-orbitals are used as basis. These atomic orbitals are then combined as in equation (2.11), and the coefficients are then the quantities that will have to be solved.

In this thesis only the Restricted (Closed Shell) Hartree Fock (RHF) method will be used. This assumes that every spatial orbital is doubly occupied by opposing spins and the spatial orbitals are restricted to be the same for both spins. However, this is not a valid assumption for large bond lengths. In, for example, molecular hydrogen at large bond lengths one electron should become localized to one atom, while the other should become localized at the other atom. In RHF this is however not the case since the electrons are assumed to be paired in the same spatial orbital. This problem can be solved by using unrestricted and open-shell Hartree-Fock methods, but these will not be used in this thesis due to the added complexity.

### 2.7.2 The Roothaan Equations

Introducing a basis set of \( K \) non-orthonormal atomic spatial orbitals allows one to derive the Roothaan equations

\[
FC = SCE \tag{2.35}
\]

where \( F \) is the Fock matrix, \( C \) is the coefficient matrix defined by its elements \( c_{ij} \) as defined in equation (2.11), \( S \) is the overlap matrix and \( E \) is a diagonal matrix containing the energies of all the molecular orbitals. The Fock matrix is defined as

\[
F_{ij} = \langle \psi_i | \hat{f} | \psi_j \rangle \tag{2.36}
\]

where \( \hat{f} \) is the Fock operator as defined in equation (2.31) and \( \psi_i \) is a particular spatial atomic orbital. The overlap matrix \( S \) arises because the basis functions are, in general, not orthonormal and is defined as

\[
S_{ij} = \langle \psi_i | \psi_j \rangle \tag{2.37}
\]

Written in this form the Hartree-Fock equations take the form of a generalized eigenvalue problem. However, the Fock matrix depends on the particular values in the coefficient matrix through the two-electron integrals \( h_{pqrs} \) and thus it cannot be solved directly. An iterative procedure, called Self-Consistent-Field is often employed to solve these equations.
In total, the procedure of solving the HF-equations is as follows:

1. Specify the geometry of the molecule and the basis set \( \{ \psi_i \} \)
2. Calculate the one-electron integrals \( h_{pq} \), the two-electron integrals \( h_{pqrs} \) and the overlap matrix \( S \)
3. Guess a particular form of the coefficient matrix \( C \), either randomly or through theoretical considerations. A starting point could be a diagonal matrix to approximate each molecular orbital by a single atomic orbital.
4. Calculate the Fock matrix with the coefficient matrix and the one- and two-electron integrals.
5. Use this calculated Fock matrix to solve the generalized eigenvalue problem of equation (2.35) and obtain an improved guess for the coefficient matrix.
6. Repeat this procedure until the Fock-matrix does not change anymore and thus a self-consistent field has been achieved.

The Fock-matrix only functionally depends on the \( N \) occupied spin-orbitals, so if these are given other molecular orbitals can easily be calculated from equation (2.35). These molecular orbitals are not occupied in the HF method, but they become important in post-Hartree-Fock methods where the assumption that the wave function is a single Slater determinant is relaxed. Other determinants where those higher energy spin-orbitals are occupied can then be included in the ground state.

### 2.8 Configuration Interaction

In the Hartree-Fock method the \( N \) lowest energy molecular orbitals are calculated using, for example, the Roothaan equations in equation (2.35). Higher energy molecular orbitals can also be calculated and it is through these higher energy molecular orbitals that one can reintroduce the electron correlations neglected in the HF-method.

To see how, consider the Hydrogen molecule in the minimal basis (4 spin-orbitals and thus also 4 molecular orbitals). In a minimal basis each hydrogen molecule contributes one spatial orbital to the basis, the total basis thus has 4 spin-orbitals. The amount of occupied molecular spin-orbitals is two and thus there also exist two unoccupied or virtual molecular spin-orbitals which have a higher energy than the occupied ones. If the two occupied spin orbitals are indexed \( i \) and \( j \) and the virtual spin-orbitals have indices \( p \) and \( q \) a number of different determinants can be written down as follows:

\[
| \Psi_{HF} \rangle = | \chi_i \chi_j \rangle, \quad | \Psi_{ij}^p \rangle = | \chi_i \chi_p \rangle, \quad | \Psi_{pq}^{ij} \rangle = | \chi_p \chi_q \rangle, \text{ etc.} \quad (2.38)
\]

The notation means that in the determinant \( | \Psi_{ij}^p \rangle \), relative to the Hartree-Fock determinant, orbital \( j \) is removed while orbital \( p \) is added. This determinant is also called a singly excited determinant. If the ground state is now a superposition of these determinants, for example, as

\[
| \Psi_0 \rangle = a \ | \Psi_{HF} \rangle + b \ | \Psi_{ij}^{pq} \rangle \quad (2.39)
\]
and an electron is measured to be in a $\chi_p$ spin-orbital, the wave function collapses to solely the $|\chi_p\chi_q\rangle$ determinant and the other electron can only reside in spin orbital $\chi_q$. This thus introduces a correlation between the electrons.

The idea of configuration interaction is then to represent the true ground state as a superposition of many of these determinants as

$$|\Psi_{CI}\rangle = c_0 |\Psi_{HF}\rangle + \sum_{i,j} c^p_{ij} |\Psi^p_{ij}\rangle + \frac{1}{4} \sum_{i,j,p,q} c^{pq}_{ij} |\Psi^{pq}_{ij}\rangle + \frac{1}{12} \sum_{i,j,k,p,q,r} c^{pqr}_{ijk} |\Psi^{pqr}_{ijk}\rangle + \ldots \quad (2.40)$$

If only determinants of single and double excitations are considered, the method is called Configuration Interaction Singles Doubles (CISD). If every possible determinant is included this is called Full Configuration Interaction (FCI) and this provides the best possible approximation to the ground state with that specific basis set. This scales very poorly however: if there are $K$ spatial orbitals (and thus $2K$ spin-orbitals) in the basis set, $N$ electrons in the system and one wishes to consider a $n$-tuple excitation there are $\binom{N}{n}$ possible ways to select $n$ orbitals to remove from the HF-determinant. Also, there are $\binom{2K-N}{n}$ possible selections of virtual orbitals. The total number of determinants in full CI is thus

$$p = \sum_{n=0}^{N} \binom{N}{n} \binom{2K-N}{n}. \quad (2.41)$$

For a molecule with 10 electrons and a basis set with 10 spatial orbitals this is already almost 200 000 different determinants and this grows even further for more complicated molecules. Still, this method is useful to test other, more tractable, methods for their accuracy and this is also how it will be used in this thesis.

## 2.9 Coupled Cluster methods

Full Configuration Interaction is intractable for larger molecules and thus alternative methods have been explored. Coupled cluster is one such method and a variant of it is also referred to as the ‘gold-standard’ of computational quantum chemistry[1]. The basic idea is that lower order excitations (e.g. a single excitation) can combine to form higher excitations (e.g. a double excitation). To represent this idea, first a $n$th order cluster operator and a total cluster operator are introduced as

$$\hat{T}_n = \frac{1}{n!} \sum_{i_1,i_2,\ldots,i_n} \sum_{a_1,a_2,\ldots,a_n} \theta^{i_1 i_2 \ldots i_n}_{a_1,a_2,\ldots,a_n} \hat{a}^\dagger_{i_1} \hat{a}^\dagger_{i_2} \ldots \hat{a}^\dagger_{i_n} \hat{a}_{a_1} \hat{a}_{a_2} \ldots \hat{a}_{a_n} \quad (2.42)$$

$$\hat{T} = \sum_{n=1}^{\infty} \hat{T}_n \quad (2.43)$$

where $\theta^{i_1 i_2 \ldots i_n}_{a_1,a_2,\ldots,a_n}$ are parameters that determine the weight of a particular cluster and $\hat{a}, \hat{a}^\dagger$ are standard fermionic creation and annihilation operators as described in section [2.6]. If only single and double excitations are considered the method is called (Unitary) Coupled Cluster Single Double ((U)CCSD) and the cluster operators take the explicit form

$$T_1 = \sum_{i,j} \theta_{ij} \hat{a}^\dagger_i \hat{a}_j \quad \text{and} \quad T_2 = \frac{1}{2} \sum_{i,j,k,l} \theta_{ijkl} \hat{a}^\dagger_i \hat{a}^\dagger_j \hat{a}_k \hat{a}_l. \quad (2.44)$$
If also triple excitations are considered, the method is called CCSDT, while if triples are considered only as a perturbation it is called CCSD(T). This last method is the aforementioned ‘golden standard’. To let these excitations combine to form higher order excitations an exponential ansatz is used as

$$|\Psi_{CC}\rangle = e^{\hat{T}} |\Psi_{HF}\rangle$$  \hspace{1cm} (2.45)

where the exponential operator is defined in terms of its power series:

$$e^{\hat{T}} = \sum_{k=0}^{\infty} \frac{1}{k!} (\hat{T})^k = 1 + \hat{T} + \frac{1}{2}(\hat{T})^2 + \frac{1}{6}(\hat{T})^3 + ...$$  \hspace{1cm} (2.46)

This series is in principle infinite but in practice it terminates after a certain order because applying an excitation that is larger than the amount of occupied spin-orbitals will yield zero. If then only single and double excitations are considered the equation above can be expanded as

$$e^{\hat{T}} = e^{\hat{T}_1+\hat{T}_2} = 1 + \hat{T}_1 + \hat{T}_2 + \frac{1}{2}(\hat{T}_1^2 + 2\hat{T}_1\hat{T}_2 + \hat{T}_2^2) + ...$$  \hspace{1cm} (2.47)

where the last three terms can then be seen to represent a double, triple, and quadruple excitation.

For an explicit example, consider the hydrogen molecule in the minimal basis set. This system has only two unoccupied spin-orbitals and for simplicity we then only consider single excitations in the cluster operator. The HF-state has two occupied spin-orbitals with indices 1 and 2 ($|\Psi_{HF}\rangle = |\chi_1\chi_2\rangle$) while the other unoccupied spin-orbitals have indices 3 and 4. The total cluster operator will then become

$$\hat{T} = \hat{T}_1 = \theta_{31}\hat{a}_3^\dagger\hat{a}_1 + \theta_{32}\hat{a}_3^\dagger\hat{a}_2 + \theta_{41}\hat{a}_4^\dagger\hat{a}_1 + \theta_{24}\hat{a}_4^\dagger\hat{a}_2.$$  \hspace{1cm} (2.48)

and the total wave function will then become

$$|\Psi_{CCSD}\rangle = |\chi_1\chi_2\rangle + \theta_{31} |\chi_3\chi_2\rangle + \theta_{32} |\chi_1\chi_3\rangle + \theta_{41} |\chi_4\chi_2\rangle + \theta_{42} |\chi_1\chi_4\rangle + \frac{1}{2}(\theta_{31}\theta_{42} - \theta_{41}\theta_{32}) |\chi_3\chi_4\rangle.$$  \hspace{1cm} (2.49)

The subtraction in the last term results from the antisymmetry of the Slater determinant.

It can be seen that the single order excitations combine to form a second order excitation, which is exactly the kind of behaviour intended with the CC-method. In this specific example the gain over FCI in the number of parameters is quite low, but for larger molecules this becomes significant.

A major problem in this theory is that the final wave function is not normalized, as can clearly be seen in equation (2.49). To circumvent this, an unitary variant has been proposed which has a slightly different version of the ansatz as

$$|\Psi_{UCC}\rangle = e^{\hat{T} - \hat{T}^\dagger} |\Psi_{HF}\rangle.$$  \hspace{1cm} (2.50)

Since unitary transformations preserve inner products, the final wave function is normalized and thus the variational principle can be used with the $\theta$ variables as variational parameters. However, this form is intractable on normal computers \[9\]. On quantum computers unitary transformations are natural operations and thus this form of coupled cluster can be explored on them. In this thesis several calculations will be done using this specific method.
3 Quantum Computing

The goal of quantum computing applied to the quantum chemistry problem is to improve upon the Hartree-Fock approximation while remaining computationally tractable. It is thus important to gain results equally good or better than post-Hartree-Fock methods while having a less severe scaling. The scaling of some conventional quantum chemistry methods are listed in table 3.1 [12]. Ideally, the quantum computer should calculate energies close to the exact FCI energies with a polynomial scaling.

<table>
<thead>
<tr>
<th>Method</th>
<th>Scaling</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hartree-Fock</td>
<td>(O(N^3))</td>
</tr>
<tr>
<td>Second order Møller–Plesset</td>
<td>(O(N^5))</td>
</tr>
<tr>
<td>CISD</td>
<td>(O(N^6))</td>
</tr>
<tr>
<td>CCSD</td>
<td>(O(N^6))</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>(O(N^7))</td>
</tr>
<tr>
<td>CCSDT</td>
<td>(O(N^8))</td>
</tr>
<tr>
<td>FCI</td>
<td>(O(N!))</td>
</tr>
</tbody>
</table>

The eigen-energies of a quantum system can be calculated in polynomial time on a quantum computer using, for example, the Quantum Phase Estimation Algorithm [13][14]. However, this requires a scalable quantum computer with error-correction codes [4] and this is not feasible on current generation technology, which is said to be in the Noisy Intermediate Scale (NISQ) regime [15]. This lead to the proposal of a variational approach to the quantum chemistry problem on a quantum computer, called the Variational Quantum Eigensolver [16]. This algorithm tries to take advantage of the relative ease with which some entangled states, which are hard to represent on classical computers, can be realized on quantum computers, while suppressing some errors due to the variational approach [9].

3.1 Basics of Quantum Computation

In Quantum Computing the individual elements of the computer are called qubits. These elements can not only be a 1 or a 0 as in a classical computer, but can be in an arbitrary superposition of these values. To represent this pictorially the Bloch Sphere is often used, shown in figure 3.1. A Pauli spin-matrix acting on a qubit then rotates the state vector around that axis.

The state of a single qubit can be characterized by only two numbers: the coefficients of the \(|0\rangle\) and \(|1\rangle\) states. For example, a qubit in the state

\[
|\Psi\rangle = \frac{1}{\sqrt{2}} (|0\rangle - |1\rangle)
\] (3.1)
can be characterized by the vector

\[
\begin{pmatrix}
c_0 \\
c_1
\end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix},
\tag{3.2}
\]

where the subscript indicates the value of the corresponding state vector. This is equivalent to writing equation (3.2) in a linear algebra fashion by taking the state vectors as basis vectors in a complex vector space:

\[
|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \text{and} \quad |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.
\tag{3.3}
\]

To extend this notation to an arbitrary number of qubits the tensor product is taken of the single qubits. For example, two qubits both in the state of equation (3.1) have a total state of

\[
|\Psi\rangle = \frac{1}{2} (|0\rangle_1 - |1\rangle_1) \otimes (|0\rangle_2 - |1\rangle_2)
\tag{3.4}
\]

where the subscript refers to the specific qubit, but this is most of the time dropped. In the linear algebra notation the tensor product is also used and the final state is then

\[
|\Psi\rangle = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{pmatrix}
\tag{3.5}
\]

where the subscripts of the last state vector refer to the binary value of the combination of those two qubits, e.g. $|0\rangle \otimes |1\rangle \equiv |01\rangle$ has a binary value of 1. The four basis vectors $\{ |00\rangle, |01\rangle, |10\rangle, |11\rangle \}$ are also said to form a computational basis and this readily extends to an arbitrary number of qubits. To completely describe a quantum system with $n$ qubits

---

**Figure 3.1:** The Bloch sphere. The state, depicted by $|\Psi\rangle$ in the figure, can be in a superposition of $|0\rangle$ and $|1\rangle$ with relative amplitudes determined by $\theta$ and with some phase factor between those two states determined by $\phi$. Image taken from [17].
one would need $2^n$ coefficients.\footnote{Since only the relative phase has physical meaning one coefficient can be eliminated and the actual number of coefficients is then $2^n - 1$. However, the length of the vector will still be $2^n$.} States on quantum computers can thus be represented by a vector in a complex vector space of dimension $2^n$. Operators acting on these states should then also have a representation of a $2^n$ by $2^n$ matrix. This means that the conjugate transpose of the matrix is its inverse. The Pauli-spin matrices are such operators acting on a single qubit and have a matrix representation of

$$
\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}.
$$

(3.6)

If multiple single-qubit operations are to be applied to several qubits at once, one should use the tensor product again. If, for example, one would want to apply a Pauli spin-matrix $\sigma_x$ to only the second qubit the total operator is

$$
\hat{O} = I \otimes \sigma_x
$$

(3.7)

where $I$ is the 2 by 2 identity matrix. This also easily extends to an arbitrary number of operations on an arbitrary number of qubits. Two-qubit gates, where two qubits are impacted together, have a matrix representation of a 4 by 4 matrix. If applying these gates to an ensemble of qubits larger than two the two-qubit gate the total operator will then have one less tensor product. As an example, the total operator for a CNOT gate applied to qubits 3 and 4 of a 4 qubit system with another $\sigma_x$ applied to the first qubit is

$$
\hat{O} = \sigma_x \otimes I \otimes \text{CNOT}.
$$

(3.8)

If the two-qubit gate is not applied to two adjacent qubits first two qubits should be swapped, and after the operation is done be swapped again.

### 3.2 Entanglement and superposition

A state is entangled if it cannot be written as simply a product state as in equation (3.4). An example of an entangled state is

$$
|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle_1 \otimes |0\rangle_2 + |1\rangle_1 \otimes |1\rangle_2) \equiv \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle).
$$

(3.9)

The electron correlation mentioned in section (2.7) can be seen to be some form of entanglement and thus entangled states can play a role in improving the HF-state. One of the gates that can introduce this entanglement is the Controlled-NOT gate (CNOT) which in the basis \{ $|0\rangle$, $|1\rangle$ \} has a matrix representation of

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

(3.10)
This gate has the effect of applying an $\sigma_x$ (which is also called a NOT gate) operator to the second qubit if the first qubit is $|1\rangle$ and applying the identity operator $I$ if it is $|0\rangle$. The first qubit is then called the control qubit and the second qubit the target qubit. To see how it entangles a state, consider the non-entangled state

$$|\Psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) \otimes |0\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |10\rangle) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$  \hspace{1cm} (3.11)

Applying the CNOT gate in equation (3.10) to this state gives the entangled state

$$|\Psi\rangle = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix} \times \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle) \hspace{1cm} (3.12)$$

The CNOT gate is in a quantum circuit diagram often represented as in figure 3.2a. Another important single-qubit gate for quantum computation is the Hadamard gate. This gate brings a state into a superposition, but is also useful for other purposes such as constructing a CNOT gate from other logic gates. Its matrix representation is

$$Hadamard = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$  \hspace{1cm} (3.13)

and its circuit symbol is shown in figure 3.2b. It also can be used to change the basis of a qubit, which will be useful in implementing coupled cluster on a quantum computer.

![CNOT gate](image1)

(a) The circuit symbol for a CNOT-gate. The upper line is the control qubit, the lower the target

![Hadamard gate](image2)

(b) The circuit symbol for the Hadamard gate.

Figure 3.2: The circuit symbols for the CNOT gate and the Hadamard gate

### 3.3 Hamiltonian simulation

The action of a Hamiltonian on a quantum system is defined by

$$|\Psi(t)\rangle = e^{-iHt} |\Psi(0)\rangle.$$  \hspace{1cm} (3.14)
If the Hamiltonian has only a single component that acts non-trivially on the system it can be used to exponentiate this term. For example, a Hamiltonian that only acts with a $\sigma_z$ operator on the third qubit (with interaction strength $\alpha$) will apply an exponentiated Pauli-matrix to the system:

$$\hat{H} = \alpha \sigma_z^{(3)} \rightarrow |\Psi_n\rangle = e^{-i \alpha t \sigma_z^{(3)}} |\Psi_0\rangle$$ \hspace{1cm} (3.15)

where the combination of $\alpha t$ determines the extent of the phase shift. The single-qubit operation of a exponentiated Pauli matrix can be shown to be equal to

$$e^{i \theta \sigma_j} = \cos(\theta) I - i \sin(\theta) \sigma_j.$$ \hspace{1cm} (3.16)

This can then be used to implement selective phase shifts and superpositions on the system, which will be useful in implementing the trial states of the VQE.

Sometimes it is interesting to calculate the time evolution of a quantum system. However, this time evolution can be very complicated for larger systems and calculating this on a classical computer is very hard. Quantum computers can do this quicker and this process is called Hamiltonian simulation. If the Hamiltonian can be written as a tensor product of Pauli-operators as

$$\hat{H} = \alpha \bigotimes_{q=1}^{n} \sigma_j^q,$$ \hspace{1cm} (3.17)

where $q$ is the number of the qubit and $\sigma_j$ is a single-qubit operator of the set $\{I, \sigma_x, \sigma_y, \sigma_z\}$, there exists a protocol to implement this Hamiltonian efficiently on a quantum computer [18]. It consists of a sort of chain of CNOT-gates to compute the parity of the system, storing the result in the last qubit, applying a phase shift to the last qubit and then reversing the chain to uncompute the parity. Qubits that have the identity operator applied to them can be ignored and do not have to participate in the circuit. If the $\sigma_x$ or $\sigma_y$ operator is applied to a qubit in the Hamiltonian the basis first has to be changed and after the computation reversed. If the Hamiltonian acts with $\sigma_x$ on qubit $k$, the Hadamard gate should applied to the qubit $k$ before and after the computation. If the Hamiltonian acts with $\sigma_y$ on qubit $k$, the operator for the basis change is

$$\tilde{R}_x \equiv Hadamard \times e^{i \frac{\pi}{4} \sigma_z}.$$ \hspace{1cm} (3.18)

After the computation the conjugate transpose should be used to reverse the basis change. An example of implementing the Hamiltonian

$$\hat{H} = \alpha \sigma_y \otimes \sigma_x \otimes \sigma_x \otimes \sigma_y$$ \hspace{1cm} (3.19)

is shown in figure [3.3]. The cost of implementing this circuit for a Hamiltonian with $p$ terms that act non-trivially and $q$ terms that are not $\sigma_z$ is then $2(p-1)$ CNOT gates and $2q + 1$ single-qubit gates.

\^3Actually the conjugate transpose should be applied, but that is equal to the Hadamard gate itself
If the Hamiltonian can only be written as a sum of tensor products of Pauli operators (terms such as in equation (3.17)) it cannot be directly implemented as in figure 3.3. The Hamiltonian can then be written as

$$\hat{H}_{\text{tot}} = \sum_i \hat{H}_i$$

(3.20)

where $\hat{H}_i$ are the single tensor product terms in the Hamiltonian. The Suzuki-Trotter expansion of order $n$ can then be used to approximate the true time evolution as

$$e^{-i\hat{H}_{\text{tot}}t} \approx \left( \prod_i e^{-i\hat{H}_i \frac{t}{n}} \right)^n.$$  

(3.21)

If the order would go to infinity, this expansion would be exact. Also, if all the terms in the Hamiltonian commute the first order expansion is already exact.

### 3.4 Mapping Creation and Annihilation operators to qubit operations

In quantum chemistry the second quantization formalism, as described in section 2.6, is often used. It is then important to be able to map the creation and annihilation operators in equations (2.25) to (2.28) to qubit operations. The two most important mappings are the Jordan-Wigner and Bravyi-Kitaev transforms.

#### 3.4.1 Jordan-Wigner transform

In quantum chemistry the Jordan-Wigner mapping has an easy physical interpretation: the occupation of a molecular orbital $k$ corresponds directly to the state of qubit $k$. For example, in a system with 2 electrons and 4 total molecular spin-orbitals the Hartree-Fock
state is simply \(|\Psi_{HF}\rangle = |1100\rangle\) (3.22)

Multiple determinants, as in section (2.8), can then also be easily expressed in qubit states. For example, a superposition of a double excited state and the HF-state is

\[|\Psi_{CI}\rangle = a|\Psi_{HF}\rangle + b|\Psi_{3412}\rangle = a|1100\rangle + b|0011\rangle.\] (3.23)

The Fermionic Creation and Annihilation (FCA) operators which satisfy the necessary anticommutation and parity relations can then be defined in terms of single-qubit gates as follows:

\[
\hat{a}^\dagger_j = (\sigma_z)_1 \otimes (\sigma_z)_2 \otimes ... \otimes (\sigma_z)_{j-1} \otimes (\sigma^+)_{j} \otimes I^{N-j}
\]

\[
\hat{a}_j = (\hat{a}^\dagger_j)^\dagger = (\sigma_z)_1 \otimes (\sigma_z)_2 \otimes ... \otimes (\sigma_z)_{j-1} \otimes (\sigma^-)_j \otimes I^{N-j}
\]

where \(\sigma^\pm = \sigma_x \mp i\sigma_y\) and \(I\) is the identity matrix \([9]\). The strings of \(\sigma_z\) ensure the proper parity and anticommutation relations. This transform can then be used to map the Hamiltonian of equation (2.29) and the cluster operator of equation (2.42) to single-qubit gates and thus apply these to any quantum computer.

A hindrance in using this kind of mapping is the high single-qubit-gate cost (weight) of the FCA operators. Every FCA operator induces a train of \(\sigma_z\) since the parity is stored non-locally while the occupation is stored locally. The weight of the FCA operators is then \(O(n)\) \([20]\). Table 3.2 contains some important combinations of FCA operators. Since \((\sigma_z)^2 = I\), some of the \(\sigma_z\) cancel out in the single and double excitation operators which reduces the weight of those operators significantly, which is useful in implementing unitary coupled cluster.

An alternative is to store the parity locally while storing the occupation number non-locally, which is called the parity-mapping \([21]\). This mapping does however not improve the weight of the FCA operators since now the occupation is stored non locally and instead of a train of \(\sigma_z\) a train of \(\sigma_x\) is used.

<table>
<thead>
<tr>
<th>Operator</th>
<th>Qubit Expression</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{n}_j = \hat{a}^\dagger_j \hat{a}_j)</td>
<td></td>
</tr>
<tr>
<td>(\hat{a}^\dagger_i \hat{a}_j, \ i &gt; j)</td>
<td></td>
</tr>
<tr>
<td>(\hat{a}^\dagger_i \hat{a}_j \hat{a}_k \hat{a}_l, \ i &gt; j &gt; k &gt; l)</td>
<td></td>
</tr>
</tbody>
</table>

\[ \begin{align*}
(\sigma^-)_j \otimes (\sigma_z)_{j+1} \otimes ... \otimes (\sigma_z)_{i-1} \otimes (\sigma^+)_i \\
(\sigma^-)_i \otimes (\sigma_z)^{l-k-1} \otimes (\sigma^-)_k \otimes (I)^{j-k-1} \otimes \\
(\sigma^+)_j \otimes (\sigma_z)^{i-j-1} \otimes (\sigma^+)_i 
\end{align*} \]

Table 3.2: A table of the most important JW transforms in quantum chemistry. A superscript is shorthand for a tensor product of that operator being applied that many times.

\(^4\)Convention here is that the first qubit represents spatial orbital 1 with spin up, qubit 2 represents spatial orbital 1 with spin down, etc. Different conventions are used, for example in \([19]\) and one should be careful with this. However, in this thesis only this specific convention is used.
3.4.2 Bravyi-Kitaev transform

The Bravyi-Kitaev (BK) transform introduces a mapping which improves on the $O(n)$ weight of the FCA operators in the Jordan-Wigner transform. It does so by exploiting a binary tree structure [20].

The idea is that every even qubit (further named only by its number $j$) stores only occupation of spin-orbital $j$. However every non-even qubit stores information about other qubits according to the following rules: If $j$ is a power of two, store the modulo 2 addition of the occupation numbers of spin-orbitals 1 to $j$. If $\text{mod}_2(j) = 1$ for a certain $k$, store the modulo 2 addition of the occupation numbers of spin-orbitals $j$ and $j - 1$, if $\text{mod}_2(j) = 2$ store the modulo 2 addition of the occupation numbers of spin-orbitals $(j - 3)$ to $j$ in qubit $j$. If $\text{mod}_2(j) = r$ for arbitrary $k$ and $r$, store in qubit $j$ the modulo 2 addition of the occupation numbers of spin-orbitals $(j - 2^r + 1)$ to $j$.

This can be visualized in a matrix which relates the Jordan-Wigner occupation number basis to the Bravyi-Kitaev basis as follows [4]

$$|j_1j_2...j_n⟩ \rightarrow |b_1b_2...b_n⟩ \quad (3.26)$$

with $b_j = \sum_i(B_r)i_j$. The matrix $B_\text{ill}$ is defined recursively as:

$$B_1 = 1 \text{ and } B_{2^r+1} = \begin{bmatrix} B_r & 0 \\ 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \ldots & 0 & 1 \\ 1 & \ldots & 1 & B_r \end{bmatrix}$$

In this way the BK mapping is only defined for an amount of qubits corresponding to a power of two, but it can be extended to arbitrary numbers of qubits [4].

Since not every qubit needs to be updated when a FCA operator is applied the weight of these operators is less than $O(n)$, with the exact value being $O(\log(n))$ [21]. The expression of the FCA operators are more involved but the procedure is to apply the $\sigma_z$ operators to all the qubits that store the occupation number of the qubit $j$ and apply $\sigma_x$ to all the qubits that store information about the parity of the qubit $j$. A detailed exposition and computation of products of these operators is presented in [21].

3.5 Qubit Reduction using locked spin-orbitals

The size of the matrices and vectors needed for quantum computer simulation on a classical computer scales as $2^n$ with the amount of qubits $n$. Systems that need 12 qubits for simulation need $4096 \times 4096$ matrices, which proved to be the maximum which the notebooks used for quantum computer simulation can handle. To reduce the number of qubits it is often useful to presume that certain molecular spin-orbitals are ‘locked’. In the Jordan-Wigner transformation this has the effect of setting the first $p$ qubits permanently to 1, where $p$ is the amount of orbitals that are locked. This has then the effect forcing the state to be

$$|\Psi⟩ = |11...1⟩ \otimes |\Psi_s⟩ \quad (3.27)$$
where \( |11 \ldots 1 \rangle \) contain \( p \) 1’s and \( \Psi_s \) is the part of the state that is actually simulated. Physically, this process has the effect of neglecting electron correlations between the \( p \) locked orbitals and the other orbitals, thus lowering the accuracy. This process can be done by OpenFermion for both the JW and BK transformation to generate a Hamiltonian with a size of \( 2^{n-p} \), which is easier to work with. The same idea can be used to 'lock-out' certain high-energy orbitals in the same way, which OpenFermion can also do.
4 The VQE algorithm

The VQE algorithm can be implemented on any quantum computer. Also, some steps can be optimized for the specific hardware used and it is thus quite adaptable \cite{9}. The VQE algorithm has four main steps \cite{4}:

1: Prepare a quantum state with variable parameters on the quantum computer.

2: Measure the expectation value of the Hamiltonian of this state on the quantum computer using an appropriate form of the Hamiltonian which corresponds to physical qubit measurements.

3: Use some kind of optimization procedure to update the variational parameters such that the quantum state with these new parameters has a lower expectation value of the Hamiltonian.

4: Repeat the above processes until the desired accuracy is reached.

A schematic version of this algorithm on an arbitrary quantum computer is shown in figure (4.1). Each of the separate steps will be considered in more detail in the coming sections.

![Figure 4.1: A schematic view of the VQE algorithm. A parameterized trial state is prepared of which the expectation value of the Hamiltonian is measured. A classical computer then runs an optimization procedure and generates new parameters. This procedure repeats until the energy has converged. $q_i$ is the $i$th qubit.](image)

4.1 Preparing a quantum state

Two distinct approaches are used to prepare a specific quantum state. The first one uses a special ansatz tailored to the hardware which can take advantage of the specific properties of the used form of quantum computer. The second one is a more theoretical approach...
which seeks to improve convergence speed by starting in a state that is closer to the true exact wave function, an example is the Unitary Coupled Cluster ansatz (UCC).

### 4.1.1 Hardware Tailored Ansatz

The Hardware Tailored Ansatz (HTA) is designed to take advantage of the specific hardware used for the quantum computer. This approach can take many different forms but only the approaches of [7] and [22] will be described.

Kandala et al. [7] used the available entangling options of their transmon quantum computer and a number of Pauli gates parameterized by the variational parameter vector $\vec{\theta}$ to prepare a trial state. These two operations alternate $d$ (called the depth of the circuit) times:

$$\left| \Psi(\vec{\theta}) \right\rangle = \left[ \prod_{j=1}^{d} \left( \prod_{q=1}^{N} \hat{U}^q,j(\vec{\theta}) \times \hat{U}_{\text{ent}} \right) \right] \times \left[ \prod_{q=1}^{N} \hat{U}^q,0(\vec{\theta}) \right] |00\ldots00\rangle$$  \hspace{1cm} (4.1)

where $\hat{U}_{\text{ent}}$ is the naturally occurring entangling gate and $\hat{U}^q,j(\vec{\theta})$ is the parameterized Pauli operation on qubit $q$ at depth $j$. This operation in equation form is

$$\hat{U}^q,j(\vec{\theta}) = e^{i\theta^q,j_1 \sigma_z} \times e^{i\theta^q,j_2 \sigma_x} \times e^{i\theta^q,j_3 \sigma_z}.$$  \hspace{1cm} (4.2)

where $\{\sigma_i\}$ are Pauli spin-matrices. These operations are done by letting the system evolve with a specific Hamiltonian, as in section 3.3. The Hamiltonian should then only involve that specific Pauli operator on that specific qubit and the phase $\theta$ is determined by the time this Hamiltonian is applied. Since applying the $\sigma_z$ operation to $|00\ldots00\rangle$ does nothing, there are $p = 3N \cdot d + 3N - N = N(3d + 2)$ independent parameters.

Colless et al [22] used a similar approach on a two-qubit system. First two single-qubit rotations are applied specified by two phases and two amplitudes. Then an entangling operation is applied which is also parameterized by a phase and a length, for a total of six independent parameters.

### 4.1.2 Unitary Coupled Cluster

The Unitary Coupled Cluster ansatz seeks to implement UCC on a quantum computer. In this thesis only Unitary Coupled Cluster Singles and Doubles (UCCSD) is considered and the ansatz takes the specific form

$$\left| \Psi_a \right\rangle = e^{\hat{T} - \hat{T}^\dagger} \left| \Psi_{\text{HF}} \right\rangle = e^{\sum_{i,j} t_{ij} \hat{a}_i^\dagger \hat{a}_j + \sum_{i,j,k,l} t_{ijkl} \hat{a}_i^\dagger \hat{a}_k^\dagger \hat{a}_l \hat{a}_l - h.c.} \left| \Psi_{\text{HF}} \right\rangle$$  \hspace{1cm} (4.3)

where $h.c.$ denotes the hermitian conjugate. Using the techniques explained in section 3.4 this operator can be mapped to single-qubit gates. As an example, the complete UCC operator in the hydrogen minimal basis problem using the Jordan-Wigner transform is

$$\hat{T} - \hat{T}^\dagger = i/2 \times \{ \theta_{31}(Y_1 Z_2 X_3 - X_1 Z_2 Y_3) + \theta_{41}(Y_1 Z_2 Z_3 X_4 - X_1 Z_2 Z_3 Y_4) + \theta_{32}(Y_2 X_3 - X_2 Y_3) + 1/4 \times \theta_{4321}(X_1 X_2 (X_3 Y_4 + Y_3 X_4) + X_1 Y_2 (Y_3 Y_4 - X_3 X_4) + Y_1 X_2 (Y_3 Y_4 - X_3 X_4) - Y_1 Y_2 (X_3 Y_4 + Y_3 X_4)) \} + \theta_{42}(Y_2 Z_3 X_4 - X_2 Z_3 Y_4) \}$$  \hspace{1cm} (4.4)
where \( \{X_j, Y_j, Z_j\} \) is shorthand for the respective Pauli spin operator applied to qubit \( j \). Since permutations of the double excitation operator are the same up to a sign (due to the anticommutativity of the FCA operators) only one of them is included. Although some \( \sigma_z \) operators cancel, as explained in section 3.4.1, this is still a lengthy expression for the simplest system. The exact scaling will be explored in Appendix A while the implementation of this total operator will be discussed in section 5.3.

4.2 Measuring the Hamiltonian

Measuring the Hamiltonian is done using the second quantization formalism of section (2.6) such that the Hamiltonian is

\[
\hat{H} = \sum_{pq} h_{pq} \hat{a}_p^\dagger \hat{a}_q + \sum_{pqrs} h_{pqrs} \hat{a}_p^\dagger \hat{a}_q^\dagger \hat{a}_r \hat{a}_s. \tag{4.5}
\]

where \( \{h_{pq}, h_{pqrs}\} \) are the integrals defined in section (2.6) and \( \{\hat{a}_j, \hat{a}_j^\dagger\} \) are the fermionic creation and annihilation operators applied to spin-orbital \( j \). These spin-orbitals should be regarded as molecular spin-orbitals, which are obtained using the Hartree-Fock method. Calculating these integrals requires one to compute the HF-energies and HF-orbitals and thus any result of the VQE that is not better than the HF-energy is not useful⁵. This Hamiltonian is then transformed using the Jordan-Wigner transform of section (3.4.1) or the Bravyi-Kitaev transform of section (3.4.2) such that it is comprised of only single-qubit gates. For example, the molecular hydrogen Hamiltonian in the minimal basis is, after the Jordan-Wigner transformation

\[
\hat{H} = -0.099 + 0.169Z_1Z_2 + 0.174Z_3Z_4 + 0.171(Z_1 + Z_2) - 0.223(Z_2 + Z_3) \\
+ 0.045(X_1Y_2Y_3Y_4 - X_1X_2X_3Y_4 + Y_1X_2X_3Y_4 - Y_1Y_2X_3X_4) \\
+ 0.166(Z_1Z_4 + Z_2Z_3) + 0.121(Z_1Z_3 + Z_2Z_4), \tag{4.6}
\]

where the same shorthand notation for the Pauli-operators was used as in equation (4.4).

Expectation Values and measurements

To get the expectation value of such an Hamiltonian on the prepared trial state it is then needed to take an expectation value of every single term in the Hamiltonian and add them together. In other words,

\[
\langle \Psi | \hat{H} | \Psi \rangle = \sum_i \langle \Psi | \hat{h}_i | \Psi \rangle = \sum_i h_i \langle \Psi | (\sigma_{i_1})^1 \otimes (\sigma_{i_2})^2 \otimes ... \otimes (\sigma_{i_N})^N | \Psi \rangle \tag{4.7}
\]

where \( \hat{h}_i \) are the partial Hamiltonian operators, \( h_i \) the coefficients of those operators and \( (\sigma_{i_j})^j \) is the Pauli operator corresponding to the one in \( \hat{h}_i \) for qubit \( j \). Pauli-operators on different qubits commute, thus

\[
\langle \Psi | (\sigma_{i_1})^1 \otimes (\sigma_{i_2})^2 | \Psi \rangle = \langle \Psi | (\sigma_{i_2})^2 | \Psi \rangle \times \langle \psi | (\sigma_{i_1})^1 | \psi \rangle \tag{4.8}
\]

⁵This is not strictly true, e.g. in \(^[7]\) atomic spin-orbitals are used. However, the HF-method has a computational cost of \( O(N^3) \) and is tractable (but inaccurate) for quite large molecules. VQE calculations which do not do better than this are thus still not useful. Another reason why MO’s are used in this thesis is because OpenFermion \(^[23]\), which will be used extensively, uses this.
and the problem is simplified to measuring the expectation value of single-qubit gates. Note that measuring \((\sigma_i^2)^1\) and then measuring \((\sigma_i^1)^2\) after that does not provide the correct answer since the first measurement already projects the state into a certain configuration. It is then necessary to prepare the same initial state twice and separately measure both operators. However, more efficient ways exist to directly measure the tensor product of Pauli operators [24]. Since they are expectation values the measurements also have to be repeated several times.

**Post Rotations and linear algebra**

Measuring different types of Pauli-operators should also be done differently. Measuring an \(I\) operator is equal to doing nothing and thus the expectation value of the first term in equation (4.6) is simply the value itself. Measuring the \(\sigma_z\) operator is requires measuring the state of that qubit. Measuring the \(\{\sigma_x, \sigma_y\}\) operators requires a change of basis such that the \(x\) or \(y\) axis of the Bloch sphere becomes equal to the \(z\) axis of the Bloch sphere. Measuring the \(\sigma_x\) thus requires a rotation around the \(y\)-axis of \(\frac{\pi}{2}\) and this is achieved by applying the gate

\[
R_y\left(\frac{\pi}{2}\right) = e^{i\frac{\pi}{4} \sigma_y} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}
\]

and then measuring the \(\sigma_z\) operator. Measuring \(\sigma_y\) is similar but with \(R_x\), which is defined with an exponential \(\sigma_x\) operator.

The tensor products in \(\hat{H}\) can also be directly calculated. In a \(n\) qubit system the Hamiltonian is then a \(2^n \times 2^n\) matrix from which the expectation value can be taken as

\[
\langle \Psi | \hat{H} | \Psi \rangle = \Psi^\dagger \hat{H} \Psi
\]

where \(\Psi\) denotes the state in the linear algebra notation of section 3.1. If the quantum computer is simulated it provides a fast way of calculating the expectation value of the Hamiltonian since one does not have to introduce the post-rotations described above. However, measurement errors are ignored this way. Since the expectation value is calculated using multiple measurements random fluctuations will cancel out, but systematic errors may still propagate. Due to the variational nature this is mitigated further, but for the measurement of the final state these errors have to be taken into account. In this thesis I will not take measurement errors into account and thus equation (4.10) is used for the expectation value of the Hamiltonian on a certain trial state.

### 4.3 Optimizing the variational parameters

There are various optimization algorithms used for the VQE. Examples include Particle Swarm Optimization [22], Nelder-Mead [25][3] and Simultaneous Perturbation Stochastic Approximation (SPSA) [7]. In this thesis only the SPSA algorithm will be used. This algorithm uses only two measurements to approximate a gradient, while still converging in higher dimensional problems and being quite robust to noise [4].
The SPSA algorithm

Let $|\Psi(\vec{\theta})\rangle$ represent the quantum state parameterized by a $p$ dimensional vector $\vec{\theta}$. For a certain iteration number $k$ with vector $\vec{\theta}_k$, sample the values

$$y^\pm_k = \left\langle \Psi(\vec{\theta}_k \pm c_k \vec{\Delta}_k) \right| \hat{H} \left| \Psi(\vec{\theta}_k \pm c_k \vec{\Delta}_k) \right\rangle,$$

(4.11)

where $c_k$ is a coefficient which decreases with increasing $k$ and $\vec{\Delta}_k$ is a randomly generated vector according to some distribution (with conditions given in [26]). The gradient is then approximated as

$$\vec{g}_k = \frac{y^+_k - y^-_k}{2c_k} \vec{\Delta}_k.$$

(4.12)

The parameter $\vec{\theta}_k$ will then be updated anti-parallel to the gradient vector as

$$\vec{\theta}_{k+1} = \vec{\theta}_k - a_k \vec{g}_k,$$

(4.13)

where $a_k$ is some number that defines the weight of that particular iteration and should decrease with iteration number. A simplistic picture of the algorithm is shown in figure 4.2.

Figure 4.2: A pictorial explanation of the SPSA algorithm. The parameters $\vec{\theta}_k$ are perturbed by a disturbance dependent on the parameter $c_k$. The energies of these new parameters values are evaluated and a gradient is calculated, after which the parameters are updated according to a weight $a_k$. Based on a figure in [7].
5 Methods

In this thesis three main areas of research will be described. The first is the optimization of the initial parameters of the Hardware Tailored Ansatz of equation (4.1). The second is the (efficient) implementation of UCC on a VQE and the third is the efficient calculation of potential surfaces using the VQE. In this section the methods for researching each of them will be described, along with a section on general methods that are applicable to all three.

5.1 General Methods

The quantum system is simulated using linear algebra, as described in section (3.1). This simulation thus involves large amounts of matrix multiplications. MATLAB was then chosen as a programming language for its optimized matrix multiplication. The Hartree-Fock and FCI calculations (to compare the VQE results with the exact results) are done with Psi4 [27]. The integrals in equation (4.5) are calculated from these HF results with OpenFermion [23]. The Jordan-Wigner and Bravyi-Kitaev transformations of annihilation and creation operators are also calculated with OpenFermion.

The SPSA optimization method discussed in section (4.3) is implemented in MATLAB with the distribution $\Delta_k$ in equation (4.11) taken to be the Rademacher distribution. This distribution randomly assigns 1 or -1 to every entry in the vector, with equal chance. Unless stated otherwise, the coefficients $c_k$ and $a_k$ are taken to be

$$c_k = \frac{c_0}{k^{0.101}} = 0.01 \frac{c_0}{k^{0.101}}, \quad a_k = \frac{a_0}{k^{0.602}}, \quad a_0 = \frac{2\pi}{5} \frac{c_0}{\langle |y^+ - y^-| \rangle} \quad (5.1)$$

where $k$ is the iteration number, $y^\pm$ is defined in equation (4.11) and the average in the last equation should be taken using a number of different $\Delta_k$ vectors. This kind of normalization ensures that the parameters are reasonably updated the first few iterations of the algorithm.

5.2 Hardware Tailored Ansatz methods

The entangler $U_{\text{ent}}$ in equation (4.1) is always taken to be a simple CNOT-chain, which is shown in figure 5.1 for the four qubit case. The total circuit of the VQE as implemented in MATLAB for four qubits is then shown in figure 5.2. The parameterized $ZXX$ rotations are done in MATLAB using tensor products and predefined exponential Pauli-matrices, as in equation (3.16). The CNOT chain is defined by tensor products of the CNOT-matrix in equation (3.10). For the circuit shown in figure 5.1 this is

$$\text{CNOT}_{\text{chain}} = (\text{CNOT} \otimes I \otimes I) \times (I \otimes \text{CNOT} \otimes I) \times (I \otimes I \otimes \text{CNOT}) \quad (5.2)$$

where $\times$ denotes matrix multiplication.
Figure 5.1: The CNOT-chain used for entangling four qubits.

Figure 5.2: The circuit for implementing the HTA for four qubits. The single-qubit blocks are the $ZXZ$ rotations. The $\hat{H}$ block represents the application of equation 4.10.
The evaluation of the Hamiltonian is done by applying equation (4.10). The initial \( \vec{\theta} \) is taken to be a random vector, with values ranging from \(-\pi\) to \(\pi\). This is something that can be optimized, which will be discussed in section 6.1.4

5.3 Unitary Coupled Cluster methods

In this section two implementations of Unitary Coupled Cluster for VQE simulation on classical computers are proposed. The first one implements the UCC ansatz exactly, as in equation (4.3), but is not a version that can be implemented on actual quantum computers. The second one is more realistic and could be implemented on quantum computers. The scaling of this approach with the size of the system is also worked out.

5.3.1 Exact implementation

The exact approach uses a script written in Python to list all the different excitation operators \( \{\hat{a}_i^\dagger \hat{a}_j, \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_k \hat{a}_l\} \). From each of them the hermitian conjugate will be subtracted, after which they are transformed to single-qubit gates using the Jordan-Wigner transform. The tensor products are then carried out, as in the case of the Hamiltonian in section (4.2) to give a matrix representation of each excitation operator. These matrices are then multiplied elementwise by their respective value of \( \theta_{ijkl} \) given by the \( \vec{\theta} \) vector, added up and exponentiated by the built in MATLAB function \textit{expm}. This gives a new matrix, which will be multiplied by the HF-state to give the trial state. This process is shown in figure (5.3) with a flowchart.

![Flowchart](image)

**Figure 5.3:** The flowchart for the exact and unrealistic implementation of UCC on a simulation of a quantum computer on a classical computer. The upper boxes describe the action that is taken while the lower boxes give the intermediate result after that action. The dashed rectangle indicates the parts that have to be redone every iteration of the SPSA algorithm.
This method is not realistic is because in general one cannot implement such a complicated exponential operator with many terms and variational parameters. However, it is useful to test the UCC method and see if it can actually produce accurate results before implementing the more realistic approach.

5.3.2 Realistic Implementation

General Approach

The realistic implementation uses a first order Trotter-Suzuki approximation, equation (3.21), of the UCC trial state, equation (4.3). It thus splits up the total cluster into parts that contain only a single tensor product of Pauli-matrices and exponentiates them separately. As an example, molecular hydrogen is considered here. The total cluster operator was given in equation (4.4). For simplicity the double excitation (parameterized by θ_{4321} in equation (4.4)) is left out and only the single excitations are considered. The total trial state operator $\hat{e}^{T-T^\dagger}$ that is desired is then

$$e^{i\frac{1}{2}[θ_{31}(Y_1Z_2X_3-X_1Z_2Y_3)+θ_{41}(Y_1Z_2Z_3X_4-X_1Z_2Z_3Y_4)+θ_{32}(Y_2X_3-X_2Y_3)+θ_{42}(Y_2Z_3X_4-X_2Z_3Y_4)]}$$ (5.3)

and the total cluster operator after the Trotter-Suzuki approximation is simply

$$e^{\hat{T}-T^\dagger} \approx e^{i\frac{1}{2}θ_{31}Y_1Z_2X_3} \times e^{-i\frac{1}{2}θ_{31}X_1Z_2Y_3} \times e^{i\frac{1}{2}θ_{41}Y_1Z_2Z_3X_4} \times e^{-i\frac{1}{2}θ_{41}X_1Z_2Z_3Y_4} \times e^{i\frac{1}{2}θ_{32}Y_2X_3} \times e^{-i\frac{1}{2}θ_{32}X_2Y_3} \times e^{i\frac{1}{2}θ_{42}Y_2Z_3X_4} \times e^{-i\frac{1}{2}θ_{42}X_2Z_3Y_4}. \; (5.4)$$

If every term in this equation commuted this would be exact, but since they do not all commute, this does not represent the exact cluster operator anymore. However, conducted simulations show that it still parametrizes the considered excitations correctly. It also still has the cluster character described in equation (2.49) and is thus suitable for the quantum chemistry problem.

One Term Approximation

It is then reasonable to ask if neglecting some terms in a specific excitation operator still produces the same excitation. If all terms of the excitation are neglected except the first (arbitrarily chosen), the excitation is an One-Term-Approximated (OTA) excitation. If every excitation of the total cluster operator in equation (4.4) becomes an OTA excitation the operator will simply become

$$e^{\hat{T}-T^\dagger} \approx e^{i\frac{1}{2}θ_{31}Y_1Z_2X_3} \times e^{i\frac{1}{2}θ_{41}Y_1Z_2Z_3X_4} \times e^{i\frac{1}{2}θ_{32}Y_2X_3} \times e^{i\frac{1}{2}θ_{42}Y_2Z_3X_4} \times e^{i\frac{1}{2}θ_{4321}X_1X_2X_3Y_4}. \; (5.5)$$

Further simulations showed that in some situations, this can work, reducing the amount of terms in the single excitations by a factor two and in the double excitations by a factor eight. The cluster character of equation (2.49) is also still exhibited after application of this further approximated operator. However, the specific approximation in equation (5.5) introduces a 4-electron term, which is not physical in that system. If only the
double excitation and one single excitation is set to be an OTA excitation the total cluster operator becomes

\[ e^{T-T^\dagger} \approx e^{\frac{i}{2} \theta_{31} Y_1 Z_2 X_3} \times e^{\frac{i}{2} \theta_{4321} X_1 X_2 X_3 Y_4} \times e^{\frac{i}{2} \theta_{41} Y_1 Z_2 Z_3 Y_4} \times e^{-\frac{i}{2} \theta_{41} X_1 Z_2 Z_3 Y_4} \times e^{\frac{i}{2} \theta_{32} Y_2 Z_3 X_4} \times e^{-\frac{i}{2} \theta_{32} X_2 Z_3 Y_4} \times e^{\frac{i}{2} \theta_{42} Y_2 Z_3 X_4} \times e^{-\frac{i}{2} \theta_{42} X_2 Z_3 Y_4}. \] (5.6)

and the final state once again only has physical terms. Other systems, such as a ten qubit system with various excitations, can work fine with every term being an OTA excitation. Clear rules that define which excitation can be approximated like this could not be found, however, errors mostly arise when OTA excitations interact with each other. Excitations which interact with each other provide the cluster character of equation (2.49). Since the cluster character is the reason why UCC works in the first place, excitations should interact and errors can thus not be avoided if every excitation becomes an OTA excitation.

This approximation was also sort of made in [5] for the calculation of the LiH groundstate using UCC. However, this was with a parameterization of only two double excitations which cannot interact with each other and thus errors are mostly suppressed.

The disadvantage of these successive approximations is that the final cluster amplitudes \( \theta_{ij(kl)} \) will not directly correspond to the actual amplitudes of the excitations in equation (4.3). However, the ground state energy is the most important quantity and thus this is a minor problem at best.

**Implementation using Hamiltonian Simulation**

The equations (5.4), (5.5) and (5.6) can be regarded as successive simulations of separate Hamiltonians. The techniques in section (3.3) can then be used to implement these excitation operators. As an example, the circuit for simulating the first two terms of equation (5.5) is shown in figure 5.4.

![Figure 5.4: The circuit for simulating the first two terms in equation (5.5). The \( R_x \) and \( R_x^\dagger \) cancel each other so they do not have to be applied, they are included for illustrative purposes. The initial state corresponds to the HF-state. The dashed parts of the circuit stay the same every iteration.](image-url)
Since $H \times H = I$, $R_x \times R_x^\dagger = I$ and $CNOT \times CNOT = I$ these operations applied successively have no effect on the state. Some terms, for example the $R_x^\dagger \times R_x$ operation in figure [5.4], thus cancel. By ordering the terms efficiently some operations can then cancel and thus less gates are needed to implement the UCC operator, but this has not been investigated further.

Implementing the total cluster operator for all possible single and double excitations after using only a single order trotterization approximation (cluster operator of the form [5.4]) has a computational cost scaling with the system size as $O(N^5)$. A proof of this result is given in appendix A. Comparing this with table 3.1, it is seen that it scales better than conventional Coupled Cluster Single Double (CCSD). However, it must be kept in mind that UCC uses quantum computing resources, which are considerably more expensive than classical computing resources. Nevertheless, it can, due to the unitary and variational nature of the method, provide better results than ordinary CCSD in polynomial time.

### Code implementation

The implementation in the code is similar to the exact approach. First a list is made of the distinct excitation operators. The hermitian conjugate is subtracted and each excitation is then transformed using the JW transform. Every excitation term is then split up into single tensor products, as in equation (5.4). Then in Python a script was written which, for every term, translates the Hamiltonian Simulation techniques into tensor products of single and two-qubit operators for implementation in MATLAB. In this step some excitations can be set to be OTA excitations. The effect of the basis-changes and CNOT gates, corresponding to the parts in a dashed rectangle in figure 5.4, do not change at each iteration and thus their matrix representations can be calculated before beginning the VQE algorithm. In the last step the variational parameters are used to calculate the $Z$ phase shift matrices. These matrices are multiplied to obtain the trotterized approximation of the cluster operator. Another flowchart of this process is shown in figure 5.5

![Figure 5.5](image)

**Figure 5.5:** The flowchart for the realistic UCC implementation. HS is Hamiltonian Simulation, VP is variational Parameters. The dashed line again represents the parts that have to be repeated every iteration.
Bravyi-Kitaev and Jordan-Wigner for UCC

The Bravyi-Kitaev transformation maps annihilation and creation operators to $O(\log(N))$ single-qubit gates, while the Jordan-Wigner transformation maps the operators to $O(N)$ single-qubit gates. Thus the Bravyi-Kitaev transformation may lead to a total gate cost that is less than the Jordan-Wigner transformation. To investigate this, OpenFermion [23] was used to transform all the distinct excitation operators with both the JW-transform and the BK-transform. After this the terms were split up, as before, and the translation into the HS-type circuit was done. The amount of CNOT-gates was then counted in Python. This yielded the results of figure 5.6. The difference between the two transforms is not as dramatic as the scaling mentioned above suggests. This is because most of the $O(N)$ single-qubit gates in the JW-transform consist of $\sigma_z$ gates, which can cancel if multiple annihilation and creation operators are combined, which is what the excitation operators are.

The JW-transform uses a occupation number basis and it is then very easy to tell what physical state qubit state corresponds to. The BK-transform uses a more involved basis where the physical state cannot be easily determined from the qubit state. The difference in CNOT-gates between the BK and JW transforms is quite low for small numbers of qubits and electrons, which are the systems studied in this thesis. Because of this, and the easier physical interpretation, the JW-transform is solely used in this thesis.

![Graphs](image)

Figure 5.6: The scaling of the realistic implementation for the BK and JW transforms plotted in one figure for a fixed number of electrons and a variable number of qubits. The black squares are the values for the JW transform while the red circles are values for the BK transform.
5.4 Potential Curve methods

The potential curve of a molecule contains the values of the ground state energies of that molecule at various values of a parameter. In the case of molecular hydrogen this parameter is the internuclear separation. Such a potential curve is shown in figure 5.7 for molecular hydrogen. Potential curves can be used to study, for example, the vibrational behaviour of the nuclei, since they are constrained to move on this potential curve.

![Potential Curve](image)

**Figure 5.7:** The molecular hydrogen potential curve. Data is calculated with Psi4 using the STO-3G basis set.

Calculating such a potential curve using the VQE is straightforward. The only difference between these potential curve points is the Hamiltonian and thus the VQE algorithm simply has to be run for every point on this potential curve using the corresponding Hamiltonian. Several techniques are proposed and tested in section 6.3 to speed up convergence of the algorithm at successive points.
6 Results

The results will be subdivided into the Hardware-Tailored Ansatz, Unitary Coupled Cluster and Potential Curve results. However, some aspects of each of them will be present in the others and some comparisons will be made between them.

6.1 Hardware Tailored Ansatz Results

The potential curves of a few molecules have been evaluated using this ansatz. As mentioned in section 4.2, the HF-orbitals and energies have to be calculated in order to set up the Hamiltonian, so if the energies of the algorithm are worse than the HF-energies the result is not of much use. For each molecule the exact (for this specific basis set) FCI energies and the HF energies are compared with the VQE energies. In the first three sections the initial values of the variational parameters $\tilde{\theta}$ are all randomly chosen.

6.1.1 The Hydrogen molecule

The potential curve for molecular hydrogen using the STO-3G basis set and the JW-transformation is shown in figure 6.1. The HTA was run at a depth of 3, with depth defined in section 4.1.1. Every point on the potential curve was sampled ten times, with the value in the graph being the lowest. The agreement between the VQE energies and the FCI energies is very good, with accuracies of $10^{-5}$ Hartree. Also, the VQE energies are considerably better than the HF energies. However, at larger bond lengths this agreement is not as good.

![Figure 6.1](image-url)

**Figure 6.1:** The ground state energies of molecular hydrogen at varying bond lengths. The green triangles represent the Hartree-Fock energy, the red circles the FCI energies and the black squares the VQE energies.
The FCI state of this problem is quite simple. For example, at 0.7 Å the FCI state is
\[
|\Psi_{\text{FCI}}\rangle = \sqrt{0.9890} |1100\rangle + \sqrt{0.0110} |0011\rangle,
\]
where the notation of the occupation basis was used. Note that this does not involve single-excitations. This is a consequence of Brillioun’s theorem, which states that singly-excited determinants have a matrix element with the HF-state of zero. They can only arise in ground states due to coupling with higher excitations and their weights in the ground state are very low.

To investigate the capability of the VQE for more interesting final states other, more complex, quantum chemistry problems have been investigated. These problems include using the more sophisticated basis set of cc-pVDZ and larger molecules such as difluoride (F₂) and Lithium-Hydride (LiH). If the cc-pVDZ basis is used for molecular hydrogen the HF calculation will yield 20 molecular spin-orbitals and thus 20 qubits are needed to implement this calculation with the VQE. To reduce the amount of qubits only the ten lowest energy molecular spin-orbitals are taken into account (section 3.5) and it becomes a ten qubit problem. The FCI state is then a complicated mixture of the HF-state and double and single excitations. Since ten qubits take quite a long time to simulate the sampling per potential curve point is reduced to four and the amount of potential curve points is reduced. The potential curve using this basis set is shown in figure 6.2.

![Figure 6.2](image)

**Figure 6.2**: The potential curve of molecular hydrogen using the cc-pVDZ. The black squares represent the lowest eigenvalue of the Hamiltonian, the green triangles the HF-energies and the red circles the VQE-calculated energies. The blue triangles represent the FCI energies of the STO-3G basis set which are included to emphasize the effect of a more sophisticated basis set.

Note that the VQE energy now was not compared with the FCI-energy but with the 'Eigenvalue' energy. This 'Eigenvalue' energy is simply the lowest eigenvalue of the Hamiltonian.
restricted to the ten lowest energy orbitals, which, because of this restriction, does not represent the FCI energy. If no restriction was applied this eigenvalue energy is equal to the FCI energy. This is a fairer measure of the quality of the VQE results, since the VQE simply cannot give a result lower than this eigenvalue.

The accuracy of the VQE is quite poor in this problem, never even reaching the HF-energy. Several reasons exist which can explain this result. Firstly, the sampling per potential curve point is reduced. The ground state is also far more complicated and thus it is harder for the VQE to reach this state. The search space becomes also very much larger due to the increased amount of qubits. The search space of the STO-3G basis set was \( \text{dim} = N(3d + 2) = 44 \) dimensional, while the search space of the cc-pVDZ basis set becomes \( \text{dim} = 110 \) dimensional. The SPSA algorithm might not be optimal for such large dimensional spaces and other algorithms might thus improve the results. The search space also contains states that are not physical in this system, such as states with > 2 electrons, which pollute the search space and makes it harder for the optimization algorithm to find the global minimum. This pollution is one of the reasons why UCC can provide more accurate results and faster convergence since it parametrizes only states that correspond to physical excitations of the HF-state. Lastly, the depth of HTA might not be high enough, but higher depths have even larger search space dimension and simulations using higher depths did not show any improvements.

6.1.2 The LiH molecule

The LiH molecule has also been investigated. The minimal basis set STO-3G would yield twelve molecular spin-orbitals and thus need twelve qubits. This was reduced to four by 'locking in' (section 3.5) the two lowest energy molecular spin-orbitals and further considering the only the four next lowest energy spin-orbitals. The final state was then

\[
|\Psi\rangle = |11\rangle \otimes |\Psi_{\text{sim}}\rangle \otimes |000000\rangle
\]

where \( |\Psi_{\text{sim}}\rangle \) is the simulated part of the state and should contain two electrons. This major reduction in qubits has a large accuracy cost, but in this context the main interest is finding out whether the VQE can find low energy states, not the exact ground states of molecules. This was already out of the question because the low-accuracy STO-3G basis set was used.

The potential curve for LiH is then shown in figure 6.3. The agreement between the VQE energies and the eigenvalue energies is again quite good. The part of the curve with bond lengths lower than 2.4Å has almost equal HF and eigenvalue energies and the VQE also reaches the same energies. The closeness of the eigenvalue and HF energies is because of the qubit reduction above. The larger the qubit reduction the more electron correlations are neglected and thus the eigenvalue of the Hamiltonian will become closer to the HF-energy.
At larger bond lengths the difference between HF-energies and eigenvalue energies is larger. This is because the HF-state requires both (active) electrons to be in identical spatial orbitals, which is not feasible at larger bond lengths where it is expected that three electrons will tend to be closer to lithium and one electron will be closer to hydrogen (as in section ??). Thus the two active electrons will occupy different spatial orbitals and the HF-energy is quite a bit higher than the eigenvalue energies. The VQE at these larger bond lengths is very close to the eigenvalue energy and is a major improvement over the HF-energy. However, since this is again a four qubit problem the state of the lowest eigenvalue cannot be very complicated and is in fact quite similar to the hydrogen molecule.

Including all the molecular spin-orbitals was also tried and results of these simulations are qualitatively similar to molecular hydrogen using the cc-pVDZ basis set. The VQE-energies are almost always worse than the HF-energies and thus not useful. Because it is then a twelve qubit problem the total time needed for simulating the VQE-algorithm also became quite prohibitive and not much further testing was done with this problem.

6.1.3 The F$_2$ molecule

The F$_2$ molecule in the minimal basis set has twenty spin-orbitals. This is far too many for simulation and thus this has to be reduced. After an iterative process it was found that locking in the first twelve molecular spin-orbitals produced an interesting test case.
for the VQE algorithm. The ground state of the obtained Hamiltonian is then
\[ |\Psi\rangle = \sqrt{0.931} |11111100\rangle + \sqrt{0.0174} |11110011\rangle + \sqrt{0.0174} |11001111\rangle + \sqrt{0.0674} |00111111\rangle. \]  
(6.3)

which is a nicely entangled state and an eight qubit system suitable for simulation. The potential curve for \( F_2 \) was then calculated using this qubit-reduced Hamiltonian and compared with the HF-energy and the eigenvalue energy, which is shown in figure 6.4.

![Figure 6.4: The \( F_2 \) potential curve using the STO-3G basis set. The black squares are the eigenvalue energies, the green triangles the HF-energies and the red circles the VQE energies.](image)

Again, the VQE does not converge nicely to the eigenvalue energies for much of the same reasons outlined in the cc-pVDZ hydrogen section. At larger bond lengths the VQE energies are better than the HF energies but this can be attributed to the unphysical HF-states which restrict the electrons to reside in the same spatial orbitals, as explained in the LiH section.

### 6.1.4 Initial state optimizing of the HTA

The poor results of the algorithm for more entangled states motivate the optimization of the initial variational parameters \( \vec{\theta} \). An obvious candidate is a \( \vec{\theta} \) that generates a HF-state since this is also the starting point for post-HF methods such as Configuration Interaction. This can be achieved by having all the angles of the rotations in the HTA be zero except the last \( \sigma_x \)-rotations on the first \( n_e \) qubits, which should have angles of \( \frac{\pi}{2} \). VQE simulations showed that the HF-state proved to be a local minimum in the parameter-space. The VQE energy obtained by running the VQE algorithm was then simply the HF-energy and thus not a useful result. This local minimum can be escaped in a few ways. First, the
SPSA parameters could be updated to sample points outside the local minimum for the approximated gradient. However, this would require the $c_k$ to be larger which would yield to less accurate gradients and thus a more noisy energy descent. Simulations showed that this did not improve upon the HF-energy.

A second approach is to modify the $\vec{\theta}$ values of HF-state by some small, random perturbation as in

$$\vec{\theta}_p = \vec{\theta}_{HF} + \alpha \vec{\Delta} \quad (6.4)$$

where $\alpha$ is a small constant that represents the weight of the perturbation and $\vec{\Delta}$ is chosen to be an uniformly random distribution with values between 0 and 1. If $\alpha$ is too large, the initial parameters will become too random and if it is too small the algorithm will become stuck in the local minimum again. The optimal value also depends on the depth of HTA ansatz. For a depth 8 HTA in the $F_2$ problem the optimal value proved to be on the order of $10^{-6}$.

The potential curve using this HF-state as initial state for $F_2$ is shown in figure 6.5. The VQE energies provide only a small improvement over the HF-energies for bond-lengths smaller than 1.6. For example, at a bond length of 1.25 Å the HF-energy is -195.9731 Hartree while the VQE energy became -195.9735 Hartree. The eigenvalue energy was -196.01465 Hartree. The VQE thus improved the HF-energy by only 0.0004 Hartree and the VQE energy did not come significantly closer to the actual eigenvalue energy. The search space around this HF-state thus seems to be quite 'flat' and the use of this initial state optimizing limited in this small-bond-length regime.

![Figure 6.5](image.png)

**Figure 6.5:** The potential curve for $F_2$ using a HF-initial state. The black squares represent the eigenvalue energies, the green triangles the HF-energies and the red circles the VQE energies using the HF initial state.
At larger bond length the VQE energy converges much better to the eigenvalue energy and is a major improvement on the HF-energy and the random parameter HTA. The sudden convergence after 1.6 Å can again be explained by the unphysical same-spatial-orbital restriction that the HF-method places on the electrons. This restriction removes the local minimum around the HF-state and the VQE can even improve on the HF energy without the small perturbation in equation (6.4). For larger bond lengths this initial parameter optimizing can thus lead to better convergence.

6.2 Unitary Coupled Cluster Results

The HTA has problems reaching highly entangled ground states, as proven by the poor results of the F₂ problem. Unitary Coupled Cluster only parametrizes states that are physical and correspond to excited states of the HF-state. The exact implementation of UCC (section 5.3.1) with all possible single and double excitations has been applied to the F₂ problem and the result is shown in figure 6.6.

![Figure 6.6: The F₂ potential for the exact UCC method. Black squares are the eigenvalue energies, red circles the VQE energies and green triangles the HF energies.](image)

The VQE converges very well to the exact eigenvalues for small and large bond lengths. This is much better than the HF-energies and a major improvement to the HTA energies of the previous sections. The realistic implementation is quite computationally expensive to simulate due to the large amounts of tensor products and matrix multiplications. A few points on this potential curve were then sampled and this converged equally well as the exact implementation. However, there are some caveats to these results. First, this F₂ molecule only possesses double excitations so the cluster effect of UCC could not be properly tested. Also, the UCC implementation is much more costly in the amount of
CNOT-gates than the HTA implementation. Implementing all single and double excitation would cost on the order of $10^3$ CNOT gates while a 8 depth HTA costs only 56 CNOT gates.

For other molecules no complete potential curve could be simulated due to time constraints, but a few points on the cc-pVDZ molecular hydrogen potential curve have been tried. There UCC showed better agreement than the HTA trial state but did not converge fully to the eigenvalue energy. This discrepancy can be explained by a few reasons: the 10-qubit UCC state takes a very long time to calculate and thus less iterations could be done than for $F_2$. More iterations could increase the accuracy but the algorithm did not seem to be improving the energy after a few hundred iterations. The larger dimensionality of the search space due to the increase in possible excitations could also be an explanation.

A few simple improvements could be made to the UCC implementation to increase the convergence and be more efficient with CNOT gates. In looking at eigenvalue ground states it became clear that certain double excitations are much less likely than others. For example, a high energy orbital and a low energy orbital rarely excite together. These further theoretical considerations could be used to lower the amount of excitations and thus lower both the amount of CNOT gates needed and the dimensionality of the search space. Also, the initial values of the variational parameters could be optimized. If ordinary coupled cluster has been run information about the importance of certain excitations could be used to initialize those excitations with higher starting values than others. Lastly, the values of the SPSA algorithm were optimized for the HTA and might not be ideal for UCC. This was found to be very much the case in simulating the realistic implementation and the values were slightly changed to better accommodate that specific problem. However, no systematic investigation on the effect of changing these parameters has taken place. Other search algorithms might also be better suited to UCC, such as the DIRECT algorithm [28].

6.3 Potential Curve optimization

In the previous results all points on the potential curve points were calculated separately and successive VQE runs had nothing to do with each other. However, the ground state of a molecule will not change very much if a small change in bond length occurs. The idea is then to use the optimal variational parameters of one point on the potential curve (PC) as a starting point for the VQE run of the next point on the potential curve. This Feed-Forward (FF) can then drastically speed up convergence since the algorithm starts already very close to the global minimum.

6.3.1 $H_2$ Feed-Forward

This Feed-Forwarding has been applied to the $H_2$ with the STO-3G basis set. The convergence of the Feed-Forward potential curve will then be compared with the standard potential curve and the result is shown in figure 6.7 for two points on the potential curve.
6.3.2 LiH Feed-Forward

Feed-Forwarding the potential curve has also been done for four qubit LiH problem and the result is shown in Figure 6.8. The results are qualitatively the same as for H$_2$, the FF potential curve converges better and faster for every bond length.

Figure 6.7: The convergence for the FF and standard potential curve for two points. The black squares are the energy errors as a function of iteration for the standard potential curve while the red circles correspond to the FF potential curve.

The FF potential curve converges faster and to a better energy than the standard potential curve at all bond lengths. The somewhat higher error initially for the second PC point is because the state changes more between the first and second PC points than between the 20th and 21st PC points.
6.3.3 \( \text{F}_2 \) Feed-Forward

The \( \text{F}_2 \) case is more interesting since both the random initial values and the HF initial values both did not converge at all bond lengths. They did however both converge well at larger bond lengths. Because of this the potential curve was run backwards, with the last PC point feeding its final state into the starting state of the previous PC point. This FF potential curve is then compared with a potential curve generated by always using the HF-state as the starting state (as in section 6.1.4) and this is shown in figure 6.9.

![Figure 6.9: The FF potential curve (blue triangles) compared with the curve generated using the HF-state as a starting state (red circles), the HF energies (green triangles) and the eigenvalue energies (black squares).](image)

The FF potential curve does show improvement over the random ansatz potential curve of figure 6.4 but it does not improve upon the potential curve with HF-states as initial states anywhere. A potential curve generated by starting at the first point and using feed forward (as in the previous two sections) showed even less good convergence. The final state after running the VQE algorithm does not necessarily converge to the actual ground state even if it converges to a value close to the ground state energy. This state might then be somewhat further from the ground state of the next potential point, which might explain the failure of this feed-forwarding in the \( \text{F}_2 \) problem.

Every point on the potential curve is calculated by choosing the minimum energy value out of a set of a few VQE runs. An improvement to the FF scheme might then be to use the HF-state as an initial state in some and use the best state of the previous PC point for some others. This then functions as a hybrid form between feed-forwarding and using the HF-state as initial state. This was not implemented due to time constraints.
7 Conclusion

The VQE and its application to quantum chemistry was analysed using simulations on a classical computer. These simulations were conducted using a combination of Python and MATLAB.

Implementation and optimization of the Hardware Tailored Ansatz as a trial state was investigated for calculation of the ground states of the \( \text{H}_2 \), \( \text{LiH} \) and \( \text{F}_2 \) molecules. The potential curves show good quantitative agreement for \( \text{H}_2 \) and a simplified form of \( \text{LiH} \). For more complex problems such as cc-pVDZ \( \text{H}_2 \) and the \( \text{F}_2 \) molecule the VQE had trouble converging to energies close to the ground state. The VQE convergence can be improved somewhat by using the Hartree-Fock state as an initial state, but since this is a local minimum a perturbation should be applied. However, this scheme improved on the HF energy by only \( 10^{-4} \) Hartree and is thus only a very minor improvement. A feed-forward approach to potential curves yielded faster and better convergence for \( \text{H}_2 \) and \( \text{LiH} \) but for \( \text{F}_2 \) this did not improve on the Hartree-Fock initial state method. However, a hybrid form between Hartree-Fock initial state and feed-forward has been proposed which could take advantage of both.

Unitary coupled cluster for trial state preparation was investigated in two different ways. An exact approach was implemented and yielded very good results for the complex \( \text{F}_2 \) problem. However, this exact approach can not be easily implemented on a quantum computer and a more realistic approach was proposed. This realistic approach uses Hamiltonian simulation techniques and a Trotter-Suzuki expansion of the UCC operator to implement an approximated form of UCC on any quantum computer. The scaling of the quantum gate cost with the system size was proven to be \( \mathcal{O}(N^5) \) and some suggestions were given to improve this scaling, including approximating the cluster operator even further and neglecting some excitations on theoretical grounds. This realistic approach yielded the same results as the exact approach for some sampled points on the \( \text{F}_2 \) potential curve. UCC thus promises to give much better results than the Hardware Tailored Ansatz for more complicated problems such as \( \text{F}_2 \), but at a severe cost in quantum gates.
8 Outlook

Optimizing the VQE is a very complex problem due to the large amount of different possibilities in not only the trial state but also the optimization algorithm. Many aspects of the VQE are not investigated in this thesis. Examples include optimizing the SPSA parameters, using a different optimization algorithm altogether and using different entanglement options for the Hardware tailored ansatz. An error analysis is also not done and this is, because it was already done for the HTA in [7], most interesting for the realistic approach to UCC.

The HTA can be optimized further by improving the HF initial state ansatz. An idea for this might be including certain excitations in the initial state, thus both escaping the local minimum around the HF state and improving the initial guess. Different entanglement options that are easier to implement than a CNOT chain can also be investigated. Higher depth HTAs using an optimization algorithm more suited to very large dimensional search spaces might also improve convergence.

The main motivation of UCC is to let lower excitations combine to form higher excitations. No problem was found that was tractable for simulating a quantum computer on a classical computer and had a ground state that would require such combinations. This would be the most interesting test for the realistic approach to UCC, but might be too hard to simulate on a classical computer. Also, the scaling of realistic UCC can be improved in various ways, as mentioned in section 5.3.2.

The Bravyi-Kitaev transformation also has better asymptotic scaling than the Jordan-Wigner transform and implementing the BK transform thus might also be an interesting subject for further research.

Lastly, the potential curve feed-forward could be, as mentioned before, combined with the HF initial state. These optimizations could also be applied to UCC.

The VQE is a long way from calculating large molecules which can not be simulated efficiently classically. It does not converge very nicely for quite simple (relative to complex systems such as proteins) problems such as cc-pVDZ H2. However, this algorithm does not impose too many restrictions on the quantum hardware, which makes it interesting to research anyway. The avenues described above can then contribute, if only a little, to the improvement of the VQE and its quest for quantum supremacy.
9 References


A  Proof of the scaling of the realistic implementation of UCC

The proof for the scaling of UCC will be given here for the realistic implementation of UCC discussed in section 5.3.2. As mentioned in section 3.3 the cost of simulating a Hamiltonian with \( p \) non-trivially acting terms is \( 2(p - 1) \) CNOT gates. The cost in single-qubit gates with \( q \) terms acting with either \( \sigma_x \) or \( \sigma_y \) is \( 2q + 1 \). From now on only the cost in CNOT gates will be considered since these are multi-qubit entangling gates and thus the hardest to implement. The proof of the scaling will be given for the JW-transform.

A.1  Single-Excitations

For a single excitation with an excitation on qubit \( q \) and an annihilation on qubit \( p \) it can easily be verified that, after the subtraction of the hermitian conjugate the operator is of the form

\[
\hat{a}_q^+ \hat{a}_p \xrightarrow{\text{JW}} \frac{i}{2} [(X_p \otimes Z_{p+1} \otimes \ldots \otimes Z_{q-1} \otimes Y_q) - (Y_p \otimes Z_{p+1} \otimes \ldots \otimes Z_{q-1} \otimes X_q)]
\]  (A.1)

Each single excitation thus introduces two terms and a \( \sigma_Z \) train between qubits \( p \) and \( q \). The CNOT cost of one of these terms is then \( 2(q - p) \). If a system is considered with \( n_e \) electrons and a basis set that needs \( n_q \) qubits the total CNOT cost of all possible single excitations is

\[
\text{Cost} = 2 \sum_{p=1}^{n_e} \sum_{q=n_e+1}^{n_q} 2(q - p) = 2n_e n_q (n_q - n_e)
\]  (A.2)

where the factor 2 arises because of the two terms in equation (A.1) and the sum was carried out with Mathematica. Since \( n_q \) is always larger than \( n_e \) this results is always positive and scales with system size as \( \mathcal{O}(N^3) \). In the molecular hydrogen system in a minimal basis there are four single excitations, two of them with 2-tensor-products terms, one with a 1-tensor-product terms, and one with 3-tensor-product terms. This gives a total of 32 terms, which is the same as the above equation predicts.

A.2  Double Excitations

Consider a double excitation with excitations on qubits \( p \) and \( q \) and annihilations on \( r \) and \( s \) with \( p > q > r > s \). One of the eight terms of the JW-transformed excitation operator (again, after subtraction with the hermitian conjugate) will then be

\[
\hat{a}_p^+ \hat{a}_q^+ \hat{a}_r \hat{a}_s \xrightarrow{\text{JW}} \frac{i}{8} X_s \otimes (Z)^{r-s-1} \otimes X_r \otimes (I)^{q-r-1} \otimes X_q \otimes (Z)^{p-q-1} \otimes Y_p
\]  (A.3)

The terms between the sets of creation and annihilation operators are \( I \) because \( Z \times Z = I \). There are then \( |r - s| + 1 + |p - q| + 1 \) terms that are acted upon with \( \{\sigma_x, \sigma_y, \sigma_z\} \) and the total cost of one such term is

\[
\text{Cost} = 2(|r - s| + |p - q| + 1).
\]  (A.4)
The total cost of all possible double excitations is achieved by summing over all possible values of \( \{p, q, r, s\} \) and divide by four since permutations of the double excitation operator are the same up to a sign. Thus the total CNOT cost of a double excitation is

\[
\text{Cost} = \frac{1}{4} \times 8 \times \sum_{r,s=1}^{n_e} \sum_{p,q=n_e+1}^{n_q} 2(|r-s| + |p-q| + 1) \\
= \frac{4}{3} n_e (n_q - n_e)(n_q n_e(n_q - n_e) + 3n_e(n_q - n_e) - n_q)
\] (A.5)

where the sum was again carried out with Mathematica. The total cost including single excitations is then

\[
\text{Cost} = \frac{2}{3} n_e (n_q - n_e)(n_q + 2n_e(n_q - n_e)(3 + n_q))
\] (A.6)

and it is seen that this cost scales as \( O(N^5) \). In figure A.1 the scaling of systems with a fixed number of electrons and a variable number of unoccupied qubits is shown. More qubits correspond to more unoccupied orbitals and thus more distinct excitation operators. The empirical values are the exact number of CNOT gates needed, counted by the Python script used for implementing the realistic UCC approach in MATLAB.

**Figure A.1:** The analytical scaling and the empirical scaling plotted in one figure for a fixed number of electrons and a variable number of qubits. The black squares are the empirical values and the red dots are the analytical values.
In the double excitations cost formula still a lot of excitations are included in the sum which are not physical, such as $\hat{a}_i^\dagger \hat{a}_i^\dagger \hat{a}_j \hat{a}_k$. At a small system size they are a relatively large portion of the total summed over excitations. However, as system size increases they will become relatively more rare in comparison with the physical excitations and thus equation (A.6) provides a better approximation. In the two electron case the analytical value is nearly exactly double that of the empirical value. This is as expected, since half of the summed over excitations do not represent a physical excitation and the double excitations are nearly all of the CNOT cost. In the figure the behaviour described above can nicely be seen.