

# Quantum Chemistry

Eesh Gupta

February 6, 2020

## Contents

<b>1</b>	<b>Introduction</b>	<b>3</b>
<b>2</b>	<b>Classical Chemistry</b>	<b>3</b>
2.1	Elementary Quantum Mechanics . . . . .	3
2.1.1	States . . . . .	3
2.1.2	Operators . . . . .	3
2.1.3	Measurement . . . . .	5
2.2	Continuous Functions . . . . .	6
2.2.1	Hamiltonian and Born Oppenheimer approximation . . .	7
2.3	Hartree Fock Theory . . . . .	7
2.3.1	Spin Orbitals . . . . .	7
2.3.2	Hartree Products . . . . .	7
2.3.3	Antisymmetry . . . . .	8
2.3.4	Variational Principle . . . . .	10
2.3.5	Setting up . . . . .	11
2.3.6	Mean Field Operator . . . . .	11
2.3.7	Hartree Fock Recipe . . . . .	13
2.3.8	Hartree Fock Failures . . . . .	13
2.4	Post Hartree Fock Methods . . . . .	14
2.4.1	Basis Sets . . . . .	14
2.4.2	Configuration Interaction . . . . .	14
2.4.3	Coupled Cluster . . . . .	15
2.4.4	Unitary Coupled Cluster . . . . .	17
2.5	Second Quantization . . . . .	18
2.5.1	Exchange Symmetry . . . . .	18
2.5.2	Analogy . . . . .	19
2.5.3	New Operators . . . . .	20
2.5.4	Anticommutation Relations . . . . .	21
2.5.5	Notational Tweaks . . . . .	23

<b>3</b>	<b>Quantum Algorithms</b>	<b>24</b>
3.1	Quantum Phase Estimation . . . . .	24
3.1.1	Procedure . . . . .	24
3.2	Variational Quantum Eigensolver . . . . .	24
3.2.1	Variational Method . . . . .	24
3.2.2	Procedure . . . . .	24
<b>4</b>	<b>Circuit Construction</b>	<b>24</b>
4.1	Quantum Encoding Methods . . . . .	24
4.1.1	Jordan Wigner . . . . .	24
4.1.2	Parity . . . . .	26
4.1.3	Bravyi Kitaev . . . . .	27
4.1.4	Application . . . . .	28
4.2	Hamiltonian Decomposition . . . . .	28
4.2.1	Trotterization . . . . .	28
4.2.2	Exponentiation . . . . .	28
<b>5</b>	<b>Variational Quantum Eigensolver</b>	<b>28</b>
5.1	State Preparation . . . . .	29
5.1.1	Reference Wavefunction . . . . .	29
5.1.2	UCCSD Wavefunction . . . . .	29
5.2	Energy Measurement . . . . .	29
5.3	Parameter Optimization . . . . .	30
5.3.1	Direct Search . . . . .	30
5.3.2	Gradient Based Methods . . . . .	31
5.4	Qiskit Implementation . . . . .	31
<b>6</b>	<b>Quantum Phase Estimation</b>	<b>31</b>
6.1	Inverse Fourier Transform . . . . .	31
6.2	Phase Analysis . . . . .	31
<b>7</b>	<b>References</b>	<b>31</b>

# 1 Introduction

$$\langle\psi|H|\psi\rangle$$

## 2 Classical Chemistry

### 2.1 Elementary Quantum Mechanics

GOAL: Introduce the Hamiltonian

In this section, we will explore the axioms of quantum mechanics which underly the foundations of the electronic structure problem. Throughout this discussion, we will use an example of a qubit to further clarify these axioms.

What is a qubit? A qubit is a quantum equivalent of a classical bit. Recall that a bit can be either 0 or 1. In a similar way, a qubit is an electron spinning either upwards or downwards along some axis. Such “upness” and “downness” of spin is similar to 1’s and 0’s of a classical bit.

#### 2.1.1 States

This notion of spinning “up” and “down” can be made more concrete with the introduction of term states. In general, we can think of state as the “status” of a qubit, whether its spinning up  $|u\rangle$  or spinning down  $|d\rangle$ .

If this were classical mechanics, our discussion would end here. Knowing the state of a particle in classical mechanics is equivalent to knowing all that there is to know about that particle. If we know that qubit is in state  $|u\rangle$ , then we know that subsequent measurements of the spin of the particle will yield the state  $|u\rangle$ . Nothing was random in the past and nothing will be random in future. The qubit was in state  $|u\rangle$  in the past and it will remain in state  $|u\rangle$  in the future.

However, in quantum mechanics, the notion of a state is more nuanced. While a classical bit can be either 0 or 1, a qubit can be  $|u\rangle$ ,  $|d\rangle$ , or a superposition of  $|u\rangle$  and  $|d\rangle$ . For example, it is possible to prepare a state  $|A\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle$  where  $\alpha_u$  and  $\alpha_d$  are the two complex numbers. We will later see that these complex amplitudes correspond to probabilities of the state  $|A\rangle$  to be either  $|u\rangle$  or  $|d\rangle$  when measured with the appropriate operator. In this sense, a quantum state only tells us how the system was prepared. It does not tell us whether the system will be in state  $|u\rangle$  or state  $|d\rangle$  in the future. These are merely possibilities based on the complex amplitudes  $\alpha_i$ .

#### 2.1.2 Operators

While states are fundamental to quantum mechanics, they are incomplete without operators. As physicist John Wheeler explained, “operators” are machines which, given an input state, produce an output state. For example, let’s say an operator  $A$  is given an input state  $|B\rangle$ . The machine  $A$  will process  $|B\rangle$  and

output some state  $|C\rangle$ . Mathematically, this will be represented as

$$A|B\rangle = |C\rangle$$

Now, there are some states for which the operator  $A$  may output the same state as the input state.

$$A|D\rangle = d|D\rangle$$

Even though  $|D\rangle$  gains a scaling factor  $d$ , the input state  $|D\rangle$  points in the same direction as the output state  $d|D\rangle$ . These vectors or states are known as eigenvectors and their associated scaling factors are known as eigenvalues. Every transformation has its own set of eigenvectors and eigenvalues. In the case of qubit, the states  $|u\rangle$  and  $|d\rangle$  are eigenstates of the spin-z operator  $\sigma_z$  which measures spin along z axis.

$$\sigma_z |u\rangle = |u\rangle$$

(1)

$$\sigma_z |d\rangle = -|d\rangle$$

(2) Now, eigenvectors and eigenvalues are important because they correspond to measurements. While we cannot access the eigenstate of a particle, we can access the eigenvalues through our measurements. For example, as observers, we might be given a qubit and not know the state of that qubit. In that scenario, we would operate on this mystery state with  $\sigma_z$  operator. If the resulting eigenvalue or measurement is 1, then the mystery state has some component in the direction of  $|u\rangle$ . And if the measurement is -1, then the mystery state has a nonzero projection along  $|d\rangle$ .

In the above scenario, we assumed that there was nothing more to a state than its spin along  $z$  direction. However, in general, states contain a lot of information which makes it hard for us to “know” those states. We can, however, operate on that state with operators like  $\sigma_z$  and the position operator  $X$ .

Now, the eigenvalues of an operator can be imaginary. But this poses a problem, because if we were to measure the position of a particle and obtain imaginary measurements, then our results are nonsense. Thus, for operators that correspond to observables or observable quantities, their eigenvalues have to be real. Mathematically, this translates to  $A = A^{\star T} = A^\dagger$  or that the operator  $A$  is its own conjugate transpose. That is, if you were to replace all the entries of  $A$  with their complex conjugates and swap the resulting entries along the main diagonal, then the resulting matrix will be same as  $A$ . To illustrate, suppose

$$A = \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \tag{1}$$

$$\begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \xrightarrow[\text{conjugate}]{\text{complex}} \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix} \xrightarrow{\text{transpose}} \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \tag{2}$$

And indeed, the eigenvalues of  $A$  are 0 and 2. Operators like  $A$  which are their own conjugate transpose and, as a result, have real eigenvalues are known as hermitian operators. Moreover, the eigenvectors of hermitian operators with distinct eigenvalues are orthogonal. Consider the eigenvectors of the hermitian operator  $\sigma_z$ :  $|u\rangle$  with the eigenvalue  $+1$  and  $|d\rangle$  with the eigenvalue  $-1$ . Then

$$\langle d|u\rangle = \langle u|d\rangle = 0.$$

In other words, neither  $|u\rangle$  nor  $|d\rangle$  contain components along the direction of the other. Also, when we take each of their inner products with themselves, we obtain

$$\langle u|u\rangle = 1$$

$$\langle d|d\rangle = 1$$

Since  $|u\rangle$  and  $|d\rangle$  are normal and orthogonal with respect to each other, these eigenvectors form an orthonormal basis.

### 2.1.3 Measurement

Since, the eigenvectors of a hermitian operator form an orthonormal basis, we can express any state in terms of those eigenvectors. Thus, states don't have to be eigenvectors of an operator to be measured. We can measure any state by expanding it in terms of the operator's eigenvectors. For example, let's say we want to measure  $\sigma_z$  on some state  $|A\rangle$ . Initially we might think this is impossible because  $|A\rangle$  is not an eigenvector of  $\sigma_z$ . However, if we expand  $|A\rangle$  in terms of  $|u\rangle$  and  $|d\rangle$ , our task becomes much easier. Let's say the expansion of  $|A\rangle$  is as follows:

$$|A\rangle = \alpha_u |u\rangle + \alpha_d |d\rangle$$

Operating on  $|A\rangle$  with  $\sigma_z$ , we obtain

$$\sigma_z |A\rangle = \alpha_u \sigma_z |u\rangle + \alpha_d \sigma_z |d\rangle$$

Using results (1) and (2), we can simplify the above equation as follows:

$$\sigma_z |A\rangle = \alpha_u |u\rangle - \alpha_d |d\rangle$$

(3) What does this equation mean? If we were to take a qubit in the state  $|A\rangle$  and use some apparatus to measure the spin in  $z$  direction, we would not get  $\alpha_u |u\rangle - \alpha_d |d\rangle$ . Instead, we would just get a reading of 1 or -1. If we were to repeat this experiment multiple times, we would see a pattern emerging:

1. The probability of obtaining 1 as measurement is  $\alpha_u \alpha_u^*$
2. The probability of obtaining -1 as measurement is  $\alpha_d \alpha_d^*$

Suppose there is some arbitrary hermitian operator  $\sigma_i$  with  $n$  eigenvectors  $|\lambda_i\rangle$  each with different eigenvalues  $\lambda_i$ . If someone were to ask for the results of 100 measurements, we would repeatedly prepare the system in a certain state

and use our apparatus to measure  $\sigma_i$ . At the end, we would give them a table of the eigenvalues and their corresponding probabilities. But this would be too cumbersome. On occasions like these, people prefer the average or the mean value of the results. To find this quantity, we can simply add the products of the eigenvalue  $\lambda_i$  and their respective probabilities  $P(\lambda_i)$  as shown below.

$$\langle \sigma \rangle = \sum_i^n \lambda_i P(\lambda_i)$$

Here,  $\langle \sigma \rangle$  is the expected value of measuring  $\sigma_z$  on some state. In our case, the expectation value of measuring  $\sigma_z$  on  $|A\rangle$  is

$$\langle \sigma_z \rangle = 1 * (\alpha_u \alpha_u^*) + -1(\alpha_d \alpha_d^*)$$

Note that this is the same result we multiplied both sides of equation (3) by  $\langle A| = \alpha_u^* \langle u| + \alpha_d^* \langle d|$ ,

$$\langle A| \sigma_z |A\rangle = (\alpha_u^* \langle u| + \alpha_d^* \langle d|)(\alpha_u |u\rangle - \alpha_d |d\rangle)$$

$$\langle A| \sigma_z |A\rangle = \alpha_u^* \alpha_u \langle u|u\rangle + -\alpha_u^* \alpha_d \langle u|d\rangle - \alpha_d^* \alpha_u \langle d|u\rangle + \alpha_d^* \alpha_d \langle d|d\rangle$$

Since  $|u\rangle$  and  $|d\rangle$  are orthogonal,  $\langle u|d\rangle = \langle d|u\rangle = 0$ . And  $|u\rangle$  and  $|d\rangle$  being normalized implies  $\langle u|u\rangle = \langle d|d\rangle = 1$ . Thus,

$$\langle A| \sigma_z |A\rangle = \alpha_u \alpha_u^* - \alpha_d^* \alpha_d$$

Hence,  $\langle A| \sigma_z |A\rangle$  and  $\langle \sigma_z \rangle$  are equivalent notations for expectation value of observable  $\sigma_z$  on the state  $|A\rangle$ . Note that  $\langle A| \sigma_z |A\rangle$  is not the value resulting from a measurement. The apparatus will only give out the eigenvalues of  $\sigma_z$  as results. However, if we take the average of all such results, we will obtain  $\langle A| \sigma_z |A\rangle$ .

## 2.2 Continuous Functions

Before, moving on to notational tweaks, let's clarify the notion of wavefunctions. In the previous section we used  $\alpha$  to represent the amplitudes corresponding to the eigenstates of  $\sigma_z$ . In other words, if we were to project  $|A\rangle$  onto  $|u\rangle$  and  $|d\rangle$ , we would get

$$\alpha_u = \langle u|A\rangle$$

$$\alpha_d = \langle d|A\rangle$$

While the dirac notation is elegant in its treatment of states and operations upon those states, it can get cumbersome for observables with infinite number of eigenvalues. For example, the position operator can have i

What is the difference between  $|\Phi\rangle$  and  $\Phi$ ? What is expectation value?

### 2.2.1 Hamiltonian and Born Oppenheimer approximation

## 2.3 Hartree Fock Theory

### 2.3.1 Spin Orbitals

In the electronic structure problem, molecular orbitals are described by wavefunctions. Each of these wavefunctions have two components: spatial and spin.

A spatial molecular orbital  $\psi(\mathbf{r})$  is a function of a position vector  $\mathbf{r}$ . Then the probability to find an electron in a small volume  $d\mathbf{r}$  surrounding  $\mathbf{r}$  is  $|\psi(\mathbf{r})|^2$ . (Complete basis/Infinite Set discussion?)

However, the position itself does not give us a full description of the orbital. Recall that molecular orbitals can store a pair of electrons such that one of them is “spin up” and the other is “spin down”. To account for spins, we introduce the following orthonormal functions of an “unspecified spin coordinate”  $\omega$ :

$$\alpha(\omega) \rightarrow \text{“spin up”}$$

$$\beta(\omega) \rightarrow \text{“spin down”}$$

(What do we mean by spin coordinate  $\omega$ ? Is it just a dummy variable in converting the discrete spin states into continuous function s, i.e. to convert  $\langle\alpha|\alpha\rangle = 1$  into  $\int d\omega \alpha(\omega) \alpha^*(\omega) = 1$ )

By multiplying a spatial orbital with any one of these spin functions, we obtain spin orbital  $\chi(\mathbf{x})$  where  $\mathbf{x}$  specifies both spatial coordinate  $\mathbf{r}$  and spin coordinate  $\omega$ .

$$\chi_i(\mathbf{x}) = \psi(\mathbf{r})\alpha(\omega)$$

$$\chi_j(\mathbf{x}) = \psi(\mathbf{r})\beta(\omega)$$

So for every spatial orbital  $\psi(\mathbf{r})$ , we have 2 spin orbitals  $\chi_i(\mathbf{x})$  and  $\chi_j(\mathbf{x})$ . Then each spin orbital can only store 1 electron. In sum, spin orbitals give us a more detailed description of molecular orbitals, specifying both the spatial and spin components.

### 2.3.2 Hartree Products

Suppose a molecular system contains  $N$  electrons and its hamiltonian  $\mathbf{H}$  can be respresented by a sum of one electron hamiltonians  $h_i$  where each  $h_i$  describes the potential and kinetic energy of the electron  $i$ . Mathematically,

$$H = \sum_i^N h_i$$

This notation is true if we ignore electron-electron respulsions. It can also be true if each one electron hamiltonian  $h_i$  contains effects of electron-electron repulsion in some average fashion. Now, let  $\{\chi_j\}$  be a set of  $N$  eigenfunctions or *eigen-orbitals* of the hamiltonians  $\{h_i\}$  such that

$$h_i \chi_j(\mathbf{x}_i) = \epsilon_j \chi_j(\mathbf{x}_i).$$

The product of these *eigen-orbitals* forms the hartree product  $\Psi_{HP}$ .

$$\Psi_{HP}(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = \chi_j(\mathbf{x}_1) \chi_l(\mathbf{x}_2) \cdots \chi_k(\mathbf{x}_N).$$

If we were to act on  $\Psi_{HP}$  with the full electronic hamiltonian  $\mathbf{H}$ , then

$$\mathbf{H}\Psi_{HP} = \sum_i h_i \Psi_{HP}$$

$$\mathbf{H}\Psi_{HP} = h_1 \chi_j(\mathbf{x}_1) \chi_l(\mathbf{x}_2) \cdots \chi_k(\mathbf{x}_N) + \dots$$

$$\mathbf{H}\Psi_{HP} = (\epsilon_i + \epsilon_j + \dots + \epsilon_k) \Psi_{HP}$$

$$\mathbf{H}\Psi_{HP} = E \Psi_{HP}$$

So the product of the eigenfunctions  $\{\chi_i\}$  of the one electron hamiltonians  $h_i$  is the eigenfunction of the full electronic hamiltonian  $\mathbf{H}$ . And the corresponding eigenvalue  $E$  of  $\Psi_{HP}$  is the sum of the eigenvalues  $\{\epsilon_j\}$ .

Then the expression

$$|\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)|^2 d\mathbf{x}_1 d\mathbf{x}_2 \cdots d\mathbf{x}_N$$

or equivalently,

$$|\chi_j(\mathbf{x}_1)|^2 d\mathbf{x}_1 \cdots |\chi_k(\mathbf{x}_N)|^2 d\mathbf{x}_N.$$

(i) represents the probability of all electrons being in a configuration such that electron 1 occupies the volume element  $d\mathbf{x}_1$  surrounding  $\mathbf{x}_1$ , electron 2 occupies the volume element  $d\mathbf{x}_2$  surrounding  $\mathbf{x}_2$ , ....

However, such interpretation of expression (i) ignores the correlation among electrons. Suppose there are 2 electrons  $e_i, e_j$  among the set of  $n$  electrons. If  $e_i$  and  $e_j$  are sufficiently close to each other, then they would instantly repel. Hence,  $e_i$  would prefer regions of space away from  $e_j$ . Then the spatial component of  $\mathbf{x}_i$  depends on the spatial component of  $\mathbf{x}_j$ . However, expression (i) ignores such correlation when it directly multiplies  $|\chi_i(\mathbf{x}_i)|^2 d\mathbf{x}_i$  with  $|\chi_j(\mathbf{x}_j)|^2 d\mathbf{x}_j$ . In doing so, it treats the motions of  $e_i$  and  $e_j$  as independent from each other. Hence, such inability to account for correlation among electrons is a significant drawback of the Hartree product representation.

(Question: Are the basis functions designed in a way so that there is a lower probability of the spatial component of  $x_j$  being close spatial component to  $x_i$ ? If that is true, then is it still true that the hartree product representation ignores correlation?)

### 2.3.3 Antisymmetry

Consider the hartree product

$$\Psi_0^{HP} = \chi_i(\mathbf{x}_1) \chi_j(\mathbf{x}_2)$$

If we were to swap the 2 electrons i.e. switch electron 2 to the spin orbital  $\chi_i$  and electron 1 to the spin orbital  $\chi_j$ , then we would get

$$\Psi_1^{HP} = \chi_j(\mathbf{x}_1) \chi_i(\mathbf{x}_2)$$



Now, electrons are indistinguishable. This means that such swapping action should not change observable values. This is analgous to swapping the 2's in the equation “ $2 + 2 = 4$ ”. The result should be 4 in either case. Hence, in our scenario, the probability density of 2 electrons occupying the 2 spin orbitals  $\chi_i$  and  $\chi_j$  before and after the swapping action should remain same. (what probability density?) Thus,

$$P(1, 2) = P(2, 1)$$

Equivalently,

$$|\Psi_0^{HP}|^2 = |\Psi_1^{HP}|^2$$

But when we unsquare both sides, we have to attach a phase as shown below

$$\Psi_0^{HP} = e^{i\Phi} \Psi_1^{HP}$$

where  $0 \leq \Phi \leq 2\pi$ . ‘It turns out that all the fundamental particles correspond to either  $\Phi = 0$ , called bosons, or  $\Phi = \pi$ , called fermions.’ (12) Hence,

$$\Psi_0^{HP} = -\Psi_1^{HP}$$

$$\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) = -\chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)$$

Combining  $\Psi_0^{HP}$  and  $\Psi_1^{HP}$  into a single state  $\Psi_2^{HP}$ , we get

$$\Psi_2^{HP} = \frac{1}{\sqrt{2}}(\chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2))$$

where  $\frac{1}{\sqrt{2}}$  is a normalization factor. For more intuition, if we made  $i = j$  then the RHS would be 0. This makes sense because according to Pauli Exclusion Principle, no two electrons can occupy the same spin orbital. Also, note that

$$\Psi_0^{HP} = -\Psi_1^{HP} = \Psi_2^{HP}$$

(is this true?)

To represent  $\Psi_2^{HP}$  in a more concise fashion, we can use a matrix representation such that

$$\Psi_2^{HP} = \frac{1}{\sqrt{2}} \det \begin{pmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) \end{pmatrix} = \frac{1}{\sqrt{2}} \chi_i(\mathbf{x}_1)\chi_j(\mathbf{x}_2) - \chi_j(\mathbf{x}_1)\chi_i(\mathbf{x}_2)$$

This is known as the slater determinant. Note that the slater determinant has antisymmetry embedded in it, which makes it very effective tool for larger systems. Also, in the Hartree Fock theory, we will be given 2K spin orbitals out of which we will select N orbitals. These N orbitals can be written in terms of a hartree product as was done in previous section. But, because of antisymmetry

principle, it is more accurate to describe it in terms of a slater determinant. Moreover, as we move on to post hartree fock methods, we will see that Hartree Fock only contains a “single” slater determinant while the other methods such as configuration interaction contain multiple slater determinants. The adjective “single” refers to the fact that we first select  $N$  orbitals and ignore the rest of  $2K-N$  unoccupied orbitals. In contrast, “multiple” slater determinants refers to choosing more than one set of  $N$  orbitals out of the  $2K$  spin orbitals to construct the determinants. This is extremely important for electron correlation.

### 2.3.4 Variational Principle

In Quantum mechanics, observables like position or momentum are represented by hermitian operators which have the property

$$U = U^\dagger$$

Now, by the spectral theorem, the eigenvalues of a spectral matrix are real. Since we want to measure ground state energy (real quantity), we can represent Hamiltonian as a hermitian matrix whose eigenvalues  $\{\lambda_i\}$  are the energies of the various excited states. Here, the excited states correspond to the eigenvectors  $\{\psi_i\}$ . Thus, we can write out the Hamiltonian  $H$  as a weighted sum of projection operators as shown below

$$H = \sum_{i=1}^N \lambda_i |\psi_i\rangle \langle \psi_i|$$

To calculate the expectation value of  $H$  or the average measured value of  $H$  with respect to an arbitrary normalized state  $|\phi\rangle$ , we would proceed as following

$$\begin{aligned} \langle \phi | H | \phi \rangle &= \langle \phi | \left( \sum_{i=1}^N \lambda_i |\psi_i\rangle \langle \psi_i| \right) | \phi \rangle \\ &= \sum_{i=1}^N \lambda_i \langle \phi | \psi_i \rangle \langle \psi_i | \phi \rangle \\ &= \sum_{i=1}^N \lambda_i |\langle \psi_i | \phi \rangle|^2 \end{aligned}$$

Since for  $1 \leq i \leq N$ ,

$$0 \leq |\langle \psi_i | \phi \rangle|^2 \leq 1$$

we can conclude that (?!?!?!?)

$$\begin{aligned} \lambda_{\min} &\leq \sum_{i=1}^N \lambda_i |\langle \psi_i | \phi \rangle|^2 \\ \lambda_{\min} &\leq \langle \phi | H | \phi \rangle = \langle H \rangle_\phi \end{aligned}$$

This is known as the variational principle as it states that the expectation value of the hamiltonian with respect to an arbitrary normalized state is an upper bound to the exact ground state or the minimum eigenvalue  $\lambda_{\min}$ . The aim of Hartree Fock method and the VQE algorithm is to produce an approximate state  $|\psi\rangle$  such that the expectation value  $\langle H \rangle_\psi$  is really close to the  $\lambda_{\min}$  or the exact ground state energy  $E_0$ .

### 2.3.5 Setting up

Suppose we have a N electron system and we are given a set of 2K spin orbitals  $\{\chi_1, \chi_2, \dots, \chi_{2K}\}$ , the goal of Hartree Fock procedure is to find the set of N spin orbitals  $\{\chi_1, \chi_2, \dots, \chi_N\}$  such that

$$|\Psi_0\rangle = |\chi_1, \chi_2, \dots, \chi_N\rangle$$

is best single determinant approximation to the ground state. In other words, the determinant  $|\Psi_0\rangle$  gives us the lowest possible energy

$$E_0 = \frac{\langle \Psi_0 | \mathbf{H} | \Psi_0 \rangle}{\langle \Psi_0 | \Psi_0 \rangle}$$

where  $\mathbf{H}$  is the full electronic hamiltonian. In the Hartree Fock theory,  $\mathbf{H}$  has 2 components: the core hamiltonian  $h(i)$  and the mean field operator  $v_{HF}$ . For the i'th electron, the core hamiltonian is just a sum of the kinetic energy of that electron and the potential energy due to its attraction to the nuclei.

$$h(i) = -\frac{1}{2}\nabla_i^2 - \sum_a \frac{Z_a}{r_{ia}}$$

On the other hand, the mean field operator is more nuanced and will require a separate section entirely.

### 2.3.6 Mean Field Operator

The mean field operator  $v_{HF}$  consists of the coulomb term  $v_{col}$  and the exchange term  $v_{exc}$ . The colomb term concerns the repulsions between electrons while the exchange term concerns the antisymmetric property of the slater determinant. For our purposes, it is sufficient to focus on the coulomb term and ignore the exchange term.

Assume that electron one is occupying the spin orbital  $\chi_a$  and we have to account for its repulsion to all the other N-1 electrons. Then, the coulomb term for electron one is defined as

$$v_{col}(1) = \sum_{b \neq a} \int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1}$$

Let's break down this coulomb potential.

- Assume electron two is occupying some spin orbital  $\chi_b$  which is different from the spin orbital  $\chi_a$  occupied by electron one.
- $\int d\mathbf{x}_2$  is the integral over all volume elements  $d\mathbf{x}_2$  occupied by electron two. In other words, this integral describes all possible positions of electron two. Given one such position or volume element  $d\mathbf{x}_2$  ,
  - $r_{12}^{-1}$  is the two electron potential or the repulsion felt by electron one due to electron two. (Doesn't the 2 electron potential depend on electron one's position as well? What are we doing about that?)
  - $|\chi_b(2)|^2$  is the probability with which electron 2 occupies that volume element  $d\mathbf{x}_2$ .
  - $|\chi_b(2)|^2 r_{12}^{-1}$  is then the weighted repulsion felt by electron 1 due to electron 2.
- So, in computing  $\int d\mathbf{x}_2 |\chi_b(2)|^2 r_{12}^{-1}$ , we are integrating or summing over weighted repulsions due to electron two for all of its positions. This results in the expected value or average repulsion electron one feels due to electron two. (this seems not very intuitive..something is fishy)
- We then assume that electron two is occupying some other spin orbital  $\chi_b$  which is different from the spin orbital it occupied earlier and  $\chi_a$ . Proceeding as above, we calculate the average repulsion electron one feels due to electron two if electron two occupies such spin orbital  $\chi_b$ .
- So the summation  $\sum_{b \neq a}$  ensures we are calculating the average repulsions due to electron two being in all occupied spin orbitals such that  $\chi_b \neq \chi_a$ . But electron two cannot simultaneously occupy all those spin orbitals.
- Note that we could have used any of the N-1 electrons for electron two. Since electrons are indistinguishable, the electron electron repulsion term is not dependent on which nametag we assigned to the N electrons. The more important point is that we are summing repulsion due to all electrons in *occupied* spin orbitals.

Note that  $v_{col}$  is sum of one electron integrals. If it were a sum of 2 electron integrals, then it would have accounted for instantaneous repulsions between electrons. However, being a sum of one electron integrals,  $v_{col}(1)$  accounts for the the repulsions electron one feels on average due to all the other electrons. To generalize, in Hartree Fock theory, instead of electrons directly interacting with each other, each electron travels through a *field* where it feels the repulsions on average or *mean* due to the other electrons. Hence,  $v_{col}$  along with the more mathematical than conceptual exchange operator  $v_{exc}$  form the *mean field* operator  $v_{HF}$ . Together with the core hamiltonian  $h$ , it forms the fock operator  $f$  for the i'th electron as

$$f(i) = h(i) + v_{HF}(i). \quad (3)$$

So for every electron, there is a different fock operator. Now, acting on the appropriate spin orbital with fock operator  $f(i)$ ,

$$f(i)\chi(\mathbf{x}_i) = \epsilon(\mathbf{x}_i).$$

This is only true if  $\chi(\mathbf{x}_i)$  is an eigenfunction of  $f(i)$ . Then the goal of the Hartree Fock recipe is to find such eigenfunctions or “eigen-orbitals” for all of the fock operators  $f(i)$ . (is it always possible to find those eigenfunctions?)

Note that to find the fock operator  $f(i)$  for any electron  $i$ , we need to first calculate its mean field operator  $v_{HF}(i)$ . Calculation of  $v_{HF}(i)$  requires us to find the spin orbitals of all the other electrons  $\{\chi_1, \chi_2, \dots, \chi_{i-1}, \chi_{i+1}, \dots, \chi_n\}$ . So to solve for the eigenvalue equation (1), we need the solutions of the other eigenvalue equations of the form (1). Thus, the Hartree Fock equations are non-linear and thus require an iterative approach— an interesting recipe!

### 2.3.7 Hartree Fock Recipe

Now that we have all the tools, we can now see the Hartree Fock recipe in action.

1. *Guess* From the given  $2K$  spin orbitals, select  $N$  spin orbitals.
2. *Mean Fields* Use these  $N$  spin orbitals to calculate mean field operators for each electron.
3. *Operators* Use the mean fields to construct the  $N$  fock operators.
4. *Evaluation* Act on those  $N$  spin orbitals with the corresponding fock operators.
5. *Check* If this action does reproduce the spin orbitals from step 2, proceed to step 6. Otherwise, use the new set of spin orbitals to repeat steps 2, 3 and 4.
6. *Self Consistency Achieved!* The slater determinant formed from these spin orbitals is the Hartree Fock ground state.

(Does Self Consistent Field method always work) In general, the larger the given set of spin orbitals, the better the approximation to the exact ground state energy. If we were given an infinite set of spin orbitals, then the resulting energy would be defined as the *Hartree Fock Limit*. Any of the Hartree Fock approximations done with finite set of spin orbitals would be slightly higher than the *Hartree Fock Limit*.

### 2.3.8 Hartree Fock Failures

Hartree Fock method gets us very close to exact ground state molecular energy, accounting for 99% of the total energy (1). However, that remaining 1% depends on movements of electrons as they interact with each other: electron correlation energy. As electrons move about the nucleus, they encounter other

electrons which tend to repel them. These “instant” repulsions contribute significantly to molecular energies as they dictate inter molecular forces such as London Dispersion (2). However, this factor was unaccounted for in the Hartree Fock picture, where the electron felt “repulsions on average” instead of those “instant” repulsions.

To account for such electron correlation, we have to abandon pigeon-holing electrons into orbitals. If we are to accept those “instant” repulsions, we have to consider the result of those effects, mainly that electrons can be shot off to higher orbitals. Molecules can then be transitioned to excited states. In these circumstances, our wavefunction no longer consists of electrons being in only their ground state orbitals (single slater determinant). Electrons now have a non-zero probability of being in singly, doubly, triply, quadruply... excited states. As a result, the post hartree fock methods expand upon the Hartree Fock wavefunction as following:

$$\Psi = a_0\Phi_{\text{HF}} + \sum_{i=1} a_i\Phi_i$$

(1)

where  $\Phi_i$  correspond to those excited slater determinants (excited configurations of electrons/ higher orbitals being occupied). This equation presents a problem: How to determine the weights  $a_i$  of those excited configurations? We will tackle this problem in the next subsection on post hartree fock methods.

## 2.4 Post Hartree Fock Methods

### 2.4.1 Basis Sets

Before moving onto methods, it is important to clarify

### 2.4.2 Configuration Interaction

To improve upon Hartree Fock method, configuration interaction presents an obvious fix by accounting for each excited state. Note that there are different ways of achieving the same excited states. For example, in  $He_2$  with 4 electrons, the hartree fock state is all the 4 electrons residing in the 2 lowest energy MOs. But the singly excited state can be any one of those electrons being excited to any one of the unoccupied MOs. This will generate a lot of possible configurations and hence a lot of slater determinants that correspond to singly excited state of the system. Moving forward, doubly excited state corresponds to any 2 of those electrons in the ground state being excited to any combination of the unoccupied MOs. This means even more configurations and hence more slater determinants. Hence, we arrive at the equation

$$\Psi_{\text{CI}} = a_0\Phi_{\text{HF}} + \sum_{\text{S}} a_s\Phi_s + \sum_{\text{D}} a_d\Phi_d + \cdots = \sum_{i=1} a_i\Phi_i$$

(1) where each summation corresponds to all possible configurations associated with that particular excitation. When we compute the lowest energy of this

system via the variational method, we have to determine the weight  $a_i$  of every possible configuration within every excited state. It is easy to see from this discussion that as we add more electrons to a system, our computations grow factorially! Hence, for large molecules and large basis sets, full CI method is computationally impractical.

To avoid such massive scaling in terms of computations, we truncate or cut off higher excited states since they are less likely to occur than the lower excited states such as the singly or doubly excited states. While full configuration interaction gives us the exact ground state energy, our truncated models (CIS and CISD) give us an approximation, at least better than the Hartree Fock energy. And the more excited states we can account for or the more computation power we can afford, the better the approximation to exact ground state energy. While truncated CI methods give us good approximates, they are not size consistent. For example if we were to evaluate the correlation energy of molecules as the separation between them increases, the correlation energy should approach 0. And if 2 molecules A and B are separated by large enough distances then they should not interact and hence the total energy of the system should be the sum of individual energies of A and B.

$$E(A + B) = E(A) + E(B)$$

(3). However, methods such as CISD tend to give a higher energy than the sum of individual energies  $E(A) + E(B)$ . For certain systems, this difference becomes absurdly large that CISD no longer serves as a good approximation method. To avert this crisis, we turn our eyes to the “gold standard” of quantum chemistry.

### 2.4.3 Coupled Cluster

In comparison to CI, truncated Coupled Cluster is both size consistent and size extensive (?!?!). To analyze coupled cluster, let's work through the equations. The excitation operator is defined as (1):

$$T = T_1 + T_2 + \dots + T_n$$

where  $T_i$  represents the set of Slater determinants corresponding to the  $i$ 'th excited state. So when  $T_i$  operator acts on the Hartree Fock state  $\Phi_0$ , the output state becomes the set of all Slater determinants or configurations of electrons in orbitals corresponding to the  $i$ 'th excited state. So,

$$T_2\Phi_0 = \sum_{i < j}^{\text{occ}} \sum_{a < b}^{\text{vir}} t_{ij}^{ab} \Phi_{ij}^{ab}$$

where the summations on the RHS represent all possible configurations where we transfer 2 electrons from occupied orbitals to unoccupied or virtual orbitals (need to make that distinction). The  $t$ 's are the relative weights of each configuration.

At this point, the introduction of the excitation operator  $T$  may look unnecessary as it is just another way to write the configuration interaction wavefunction.

$$\Psi_{CI} = (1 + T)\Phi_0 = \Phi_0 + T_1\Phi_0 + T_2\Phi_0 + \dots$$

However, the coupled cluster wavefunction differs from the configuration interaction wavefunction in one important way: it uses an exponential operator:

$$\Psi_{CC} = e^T \Phi_0$$

When we Taylor-expand the exponential operator, we obtain

$$e^T = 1 + T + T^2 + \dots + T^n$$

$$e^T = 1 + (T_1 + T_2 + \dots + T_n) + (T_1 + T_2 + \dots + T_n)^2 + \dots + (T_1 + T_2 + \dots + T_n)^n$$

Regrouping in terms of excitations,

$$e^T = 1 + (T_1) + (T_2 + T_1^2) + (T_3 + 2T_1T_2 + T_1^3) + \dots$$

The first term on RHS is just 1, corresponding to the Hartree Fock state. The second term corresponds to the first excited state. So far so good. The third term should correspond to the second excited state. But it contains another mysterious term  $T_1^2$ . This corresponds to a product of 2 singly excited determinants, which still result in the doubly excited state. The essential difference between  $T_2$  and the mixed term  $T_1^2$  is that the former accounts for correlation between the 2 electrons while the latter term treats each electron independently. Moving onto the triply excited state term, we see 2 mixed terms,  $2T_1T_2$  and  $T_1^3$ . While both achieve the triply excited state, they do so by accounting for different amounts of correlation. Where  $2T_1T_2$  recovers at least some correlation with the  $T_2$  term,  $T_1^3$  treats electron motion as uncorrelated entirely.

So the mixed terms in our regrouped equation don't look at the whole set of electrons. They just focus on a small subset or cluster of electrons. The product of these cluster excitations achieve the same net excitation as the non-mixed term. With this exponential operator then, we can solve the schrodinger equation as

$$\mathbf{H}e^T \Phi_0 = E_{CC}e^T \Phi_0$$

$$e^{-T}\mathbf{H}e^T \Phi_0 = E_{CC}\Phi_0$$

We can modify the above equation to obtain a series of nonlinear equations by projecting excited states  $|\Phi_i^a\rangle, |\Phi_{ij}^{ab}\rangle, \dots$  onto  $\Phi_0$  as shown below

$$\langle \Phi_i^a | e^{-T} \mathbf{H} e^T | \Phi_0 \rangle = 0$$

$$\langle \Phi_{ij}^{ab} | e^{-T} \mathbf{H} e^T | \Phi_0 \rangle = 0$$

which we would have to solve variationally to determine the weights  $t$ 's of the different configurations that achieve the lowest ground state energy. Note



that above equations are equated to 0 as excited states and ground states are orthogonal i.e.  $\langle \Phi_i^a | \Phi_0 \rangle = 0$

As with configuration interaction, if the basis size is infinite, Coupled Cluster equations will force us to account for infinite number of excitations which is impractical with current computing technology. So as we did with CI, we truncate or cut off higher excitation operators to get a good approximation. To illustrate, consider the CCSD method which only accounts of singly and doubly excited state operators:

$$T = T_1 + T_2$$

In CISD, we would ignore the triply, quadruply and ... excited states. However, because of the exponential term and Taylor expansion, CCSD would recover some correlation from the higher excited states as shown below:

$$e^T = 1 + (T_2 + T_1^2) + (2T_1T_2 + T_1^3) + (T_2^2 + 3T_2T_1^2 + T_1^4) + \dots$$

The mixed terms or product of lower excited states serve as good approximations to higher excited states. Thus CCSD recovers more correlation energy than CISD, making CCSD preferable for larger systems. And due to this property, truncated coupled cluster methods are size consistent i.e correlation energy converges to 0 as we increase separation between some molecules.

#### 2.4.4 Unitary Coupled Cluster

There are 2 problems with the coupled cluster method

1.  $e^{-T}\mathbf{H}e^T$  is non hermitian because  $e^T$  isn't unitary. Unitary operators are important because they preserve norm of states. So when we try to calculate the energy  $E_a$  using the equation:

$$\langle \Phi(\theta) | e^{-T}\mathbf{H}e^T | \Phi(\theta) \rangle = E_a$$

where  $\Phi(\theta)$  is our approximation for ground state,  $E_a$  may or may not be an upper bound to the exact ground state  $E_0$ . Hence lack of unitarity implies that we can't use the variational principle (4).

2. The expansion of  $e^{-T}\mathbf{H}e^T$  converges only if we assume a single reference state. This assumption implies that we only consider the excitations from the single determinant  $\Phi_0$  or the Hartree Fock state to the excited states (1). In this case, the Hartree Fock state serves as the single reference state. However, to get a more balanced and exact correlation (11), one must also account for excitations from excited states to higher excited states. Here, every time we add excited state configurations, they will serve as reference states to generate higher excited states. Known as the Multireference Coupled Cluster, this method is especially important for strongly correlated systems where the excited states have similar energies to the ground state energy. The HF state becomes then a poor reference state as it has very small overlap with the ground state (4).

To circumvent these shortcomings of Coupled Cluster Theory, we use its unitary variant Unitary Coupled Cluster. Here we introduce the deexcitation operator in the exponential.

$$\Psi = e^{T-T^\dagger} \Phi_0$$

Here  $\Psi$  is called the Unitary Coupled Cluster state or ansatz. While the expansion of  $U = e^{T-T^\dagger}$  is infinite and hence preparation of UCC ansatz is intractable for classical computers. However, the UCC ansatz can be efficiently prepared on a quantum device in polynomial time (12).

## 2.5 Second Quantization

The notations we have used in the previous sections are termed as first quantization. In this section, we will explore second quantization and its improvement upon those notations. Note that “orbitals” refer to spin orbitals that can either be empty or hold 1 electron.

### 2.5.1 Exchange Symmetry

Suppose a hypothetical atom has the orbitals or discrete energy levels indexed as following  $A, B, C, D \dots$ . Also suppose the system only contains one electron in the state

$$|\Psi\rangle_1 = |A\rangle_1 + |B\rangle_1$$

where electron 1 is simultaneously occupying orbitals A and B. So far, our notation looks elegant and easy to interpret. But when we account for the second electron, let's say in the state  $|\Psi\rangle_2 = |C\rangle_2$  occupying orbital C, our notation for the state of whole system  $|\Phi\rangle$  gets messy. You might assume it will look like this:

$$|\Phi\rangle = (|A\rangle_1 + |B\rangle_1) \otimes |C\rangle_2$$

where electron 1 is occupying the state  $|A\rangle + |B\rangle$  while electron 2 is occupying the state  $|C\rangle$ . However, according to exchange symmetry, electron 1 and electron 2 are indistinguishable. Hence, there is an equal probability that electron 1 is occupying state  $|C\rangle$  and electron 2 is occupying the state  $|A\rangle + |B\rangle$ . In other words, if an observer is looking at a quantum system and observes some probability density  $P(1, 2)$  after we swap the electrons 1 and 2, the probability density should stay the same.

$$P(1, 2) = P(2, 1)$$

Equivalently,

$$||\Psi(1, 2)\rangle|^2 = ||\Psi(2, 1)\rangle|^2$$

But when we unsquare both sides, we have to attach a phase as shown below

$$|\Psi(1, 2)\rangle = e^{i\Phi} |\Psi(2, 1)\rangle$$

where  $0 \leq \Phi \leq 2\pi$ . ‘It turns out that all the fundamental particles correspond to either  $\Phi = 0$ , called bosons, or  $\Phi = \pi$ , called fermions.’(12) Hence we obtain the antisymmetry principle for fermions which is

$$|Psi(1, 2)\rangle = -|Psi(2, 1)\rangle$$

Coming back to our case, the combined state of the system will then be

$$|\Phi\rangle = (|A\rangle_1 + |B\rangle_1) \otimes |C\rangle_2 - |C\rangle_1 \otimes (|A\rangle_2 + |B\rangle_2)$$

which can be represented as a determinant of a matrix as shown below:

$$\det \begin{pmatrix} |A\rangle_1 + |B\rangle_1 & |C\rangle_1 \\ |A\rangle_2 + |B\rangle_2 & |C\rangle_2 \end{pmatrix}$$

The determinant of the above matrix is known as the slater determinant which was previously denoted as some configuration of the system. It is easy to see how adding more orbitals and more electrons would make writing such slater determinants really painful. This problem stems from our insistence with assigning an orbital to each electron. In the following section, we will explore Second Quantization where we will assign electrons to orbitals, reducing our writing payload.

### 2.5.2 Analogy

Suppose the state of the some system is an oscillating violin string. The violin string To represent that wave, using the fourier series formulas, we deconstruct the complicated wave into sum of simpler waves called vibrational modes. Each vibrational mode does not contribute equally to the complicated wave. Some modes contribute more than others, represented by the weight or amplitude of the mode  $a_n$ . So, we have

$$\text{complicated wave} = \sum_{\text{mode } n}^{\infty} a_n * \text{mode}_n$$

Insert GRAPHIC of wave decomposition.

In this analogy, each mode represents a particular orbital or state and the amplitude refers to the number of particles occupying that state. So, each mode  $\text{mode}_n$  or orbital is assigned an amplitude  $a_n$  or number of electrons. In contrast to first quantization, we don’t have to index our orbitals with which electron is occupying that orbital. Note that this indexing stage requires us to id each electron as electron 1, electron 2, etc. However, in attaching these name tags, we ran into issues with exchange symmetry as we were trying to distinguish what were indistinguishable. The result of applying this symmetry was more terms added to the equation.

However, in the notation under second quantization, we avoid indexing electrons or assigning specific electrons to orbitals. Instead, we tell the piper that

“this orbital exists and there are  $x$  many electrons occupying that orbital”. We don’t have to give any more information about which specific electrons belong to that orbital as we would be then assuming that electrons are distinguishable, running into the the trap set by exchange symmetry.

Moreover, in first quantization, we sum over electrons or unit amplitudes instead of the modes. According to analogy, this is nonsensical as we are treating unit amplitude or constants as separate entities. In second quantization, however, we treat the modes or orbitals as separate entities. In more elaborate language, the atom is a field, made up of different vibrational modes. Particles are then excitations or the “strength” of the vibrational modes in the overall field.

To show the underlying simplicity of the approach shown in the analogy, we turn back to our example from before. Note that each state will now be represented as

$$|n_a\rangle_A |n_b\rangle_B |n_c\rangle_C |n_d\rangle_D$$

where each  $n_i$  denotes the number of particles in the corresponding orbital. So from previous discussion,  $\Psi_1$  becomes

$$|\Psi_1\rangle = |1\rangle_A |0\rangle_B |0\rangle_C |0\rangle_D + |0\rangle_A |1\rangle_B |0\rangle_C |0\rangle_D$$

or for sake of simplicity

$$|\Psi_1\rangle = |1000\rangle + |0100\rangle$$

and  $\Psi_2$  becomes

$$|\Psi_2\rangle = |0\rangle_A |0\rangle_B |1\rangle_C |0\rangle_D$$

or

$$|\Psi_2\rangle = |0010\rangle$$

Hence the combined state is

$$|\Phi\rangle = (|1\rangle_A |0\rangle_B |0\rangle_C |0\rangle_D + |0\rangle_A |1\rangle_B |0\rangle_C |0\rangle_D) \otimes (|0\rangle_A |0\rangle_B |1\rangle_C |0\rangle_D)$$

which is equivalent to

$$|\Phi\rangle = (|1000\rangle + |0100\rangle) \otimes (|0010\rangle)$$

$$|\Phi\rangle = |1010\rangle + |0110\rangle$$

For larger systems, the simplification is even more dramatic. In the following subsection, we introduce creation and annihilation operators which are a special feature of second quantization.

### 2.5.3 New Operators

To explain the new operators and their commutation/anticommutation relations, we first introduce the vacuum state  $|\rangle$ . The vacuum state is a quantum state of length 1.

$$\langle|\rangle = 1$$

This state describes a system with no electrons. Yet even so,  $|\rangle \neq 0$ . If the vacuum state were the 0 vector, then it would be impossible for us to add electrons to the vacuum state. This is because action of any operator on the 0 vector reproduces the 0 vector. On the other hand, if we act on the vacuum state with a creation operator  $a_i^\dagger$ , we obtain

$$a_i^\dagger |\rangle = |\chi_i\rangle$$

Simply, we have added an electron to the spin orbital  $\chi_i$ . On the other hand, if we act on  $|\chi_i\rangle$  with the annihilation operator  $a_i$ , we get the vacuum state.

$$a_i |\chi_i\rangle = |\rangle$$

As the name suggests, we remove or annihilate that electron in spin orbital  $\chi_i$  to get a vacuum state i.e. a state with no electrons. Why is the creation operator  $a_i^\dagger$  notated as the adjoint of the annihilation operator  $a_i$ ? If we take the adjoint of the equation

$$|\chi_i\rangle = a_i^\dagger |\rangle,$$

we get

$$\begin{aligned}\langle\chi_i| &= \langle| (a_i^\dagger)^\dagger \\ \langle\chi_i| &= \langle| a_i\end{aligned}$$

Multiplying both sides with  $|\chi_i\rangle$ ,

$$\langle\chi_i|\chi_i\rangle = \langle| a_i |\chi_i\rangle$$

$$\langle\chi_i|\chi_i\rangle = \langle|$$

Since the states  $|\chi_i\rangle$  and  $|\rangle$  are normalized,

$$1 = 1$$

Therefore, we have verified that the creation operator  $a_i^\dagger$  is indeed the adjoint of the annihilation operator  $a_i$ .

#### 2.5.4 Anticommutation Relations

While second quantization avoids the use of Slater determinants, it cannot ignore antisymmetry principle. Even though we avoid nametagging electrons, we can still swap electrons. To illustrate, consider the state . Swapping the spin orbitals  $\chi_i$  and  $\chi_k$ , we obtain  $|\chi_i, \dots, \chi_k, \dots, \chi_l\rangle$ . This reordering corresponds to swapping the electrons (?!?!), and thus

$$|\chi_k, \dots, \chi_i, \dots, \chi_l\rangle = -|\chi_i, \dots, \chi_k, \dots, \chi_l\rangle.$$

Suppose we were to create 2 orbitals  $\chi_p$  and  $\chi_r$  in the state  $|\chi_k, \dots, \chi_l\rangle$  i.e. add 2 electrons to the unoccupied spin orbitals  $\chi_p$  and  $\chi_r$ . Does the order in

which we create these spin orbitals matter? Let's answer this question by first creating  $\chi_p$  and then  $\chi_r$ .

$$\begin{aligned} a_r^\dagger a_p^\dagger |\chi_k, \dots, \chi_l\rangle &= a_r^\dagger |\chi_p \chi_k, \dots, \chi_l\rangle \\ &= |\chi_r \chi_p \chi_k, \dots, \chi_l\rangle \end{aligned}$$

On the other hand, if we were to create  $\chi_r$  first and then  $\chi_p$ ,

$$\begin{aligned} a_p^\dagger a_r^\dagger |\chi_k, \dots, \chi_l\rangle &= a_p^\dagger |\chi_r \chi_k, \dots, \chi_l\rangle \\ &= |\chi_p \chi_r \chi_k, \dots, \chi_l\rangle \\ &= -|\chi_r \chi_p \chi_k, \dots, \chi_l\rangle \end{aligned}$$

So the order in which we create 2 spin orbitals does matter. To further formalize this notion, we first note from the above equations that,

$$(a_r^\dagger a_p^\dagger + a_p^\dagger a_r^\dagger) |\chi_k, \dots, \chi_l\rangle = 0$$

Since  $|\chi_k, \dots, \chi_l\rangle$  is an arbitrary state that does not have to equal the 0 vector,

$$a_r^\dagger a_p^\dagger + a_p^\dagger a_r^\dagger = 0$$

Thus,

$$\{a_p^\dagger, a_r^\dagger\} = 0$$

Similarly we can show that the order in which we destroy 2 spin orbitals matters. However, in this scenario, the state  $|\chi_k, \chi_j, \dots, \chi_p, \dots, \chi_r, \dots, \chi_l\rangle$  already contains the to-be-removed spin orbitals  $\chi_p$  and  $\chi_r$ . Also, note that the annihilation operator  $a_i$  can only act on some state  $|\chi_k, \dots, \chi_i, \dots, \chi_l\rangle$  if  $\chi_i$  is the leftmost spin orbital in the state. So, we have to swap spin orbitals  $\chi_i$  and  $\chi_k$  in order to act on that state with  $a_i$ . Now, coming back to our discussion of removing  $\chi_p$  and  $\chi_r$ ,

$$\begin{aligned} (a_r a_p + a_p a_r) |\chi_k, \chi_j, \dots, \chi_p, \dots, \chi_r, \dots, \chi_l\rangle \\ &= a_r a_p |\chi_k, \chi_j, \dots, \chi_p, \dots, \chi_r, \dots, \chi_l\rangle + a_p a_r |\chi_k, \chi_j, \dots, \chi_p, \dots, \chi_r, \dots, \chi_l\rangle \\ &= -a_r a_p |\chi_p, \chi_j, \dots, \chi_k, \dots, \chi_r, \dots, \chi_l\rangle - a_p a_r |\chi_r, \chi_j, \dots, \chi_p, \dots, \chi_k, \dots, \chi_l\rangle \\ &= -a_r |\chi_j, \dots, \chi_k, \dots, \chi_r, \dots, \chi_l\rangle - a_p |\chi_j, \dots, \chi_p, \dots, \chi_k, \dots, \chi_l\rangle \\ &= a_r |\chi_r, \dots, \chi_k, \dots, \chi_j, \dots, \chi_l\rangle + a_p |\chi_p, \dots, \chi_j, \dots, \chi_k, \dots, \chi_l\rangle \\ &= |\dots, \chi_k, \dots, \chi_j, \dots, \chi_l\rangle + |\dots, \chi_j, \dots, \chi_k, \dots, \chi_l\rangle \\ &= |\dots, \chi_k, \dots, \chi_j, \dots, \chi_l\rangle - |\dots, \chi_k, \dots, \chi_j, \dots, \chi_l\rangle \\ &= 0 \end{aligned}$$

Since

$$(a_r a_p + a_p a_r) |\chi_k, \chi_j, \dots, \chi_p, \dots, \chi_r, \dots, \chi_l\rangle = 0$$

and  $|\chi_k, \chi_j, \dots, \chi_p, \dots, \chi_r, \dots, \chi_l\rangle$  is an some random state that may or may not be equal to 0 vector,

$$a_r a_p + a_p a_r = 0$$

Thus,

$$\{a_p, a_r\} = 0$$

Suppose you notice that the kth orbital of some molecule is unoccupied and you want to add an electron to that orbital. Right now, the state is  $|0\rangle$  (vaccum state). So what you do is open your second quantization toolkit and pull out the creation operator  $a_k^\dagger$  and apply to the vaccum state as shown:

$$|k\rangle = a_k^\dagger |0\rangle$$

By applying that creation operator, you have changed the status of the kth orbital from being unoccupied (being vaccum state) to occupied. To undo this change, we use the annihilation operator as following:

$$|0\rangle = a_k |k\rangle$$

We will use these tools to construct the hamiltonian for the system. Before we dive into the hamiltonian jungle, its important to clarify the nature of orbitals. In the conventional sense, we refer to orbitals as a region within a subshell occupied by 2 electrons with opposite spins. From now on, these orbitals will be referred to as spatial orbitals. These spatial orbitals will each contain 2 spin orbitals which will only occupy 1 electron. This will make representing spin orbitals in terms of quantum states easier as

$$|0\rangle \implies \text{unoccupied}$$

$$|1\rangle \implies \text{occupied}$$

This allows us to use binary representation where each digit corresponds to an orbital. This is utilized in quantum fourier transform and quantum phase estimation algorithms.

### 2.5.5 Notational Tweaks

To see the effects of second quantization notation on our, we first turn to the Hartree Fock state of a n electron system:

$$\Phi_0 = a_n^\dagger a_{n-1}^\dagger \dots a_1^\dagger |0\rangle$$

In the above equation, we are transforming the vaccum state to Hartree Fock state to the first n spin orbitals of the molecule. Turning over to the hamiltonian, we obtain the following equation

$$\mathbf{H} = \sum_{i,j}^{N_{orb}} h_{ij} a_i^\dagger a_j + \frac{1}{2} \sum_{i,j,k,l}^{N_{orb}} h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l$$

(1) To make sense of this notation, let's consider ??? why are there 2 summations !?! Shouldn't there be only 1 over all the spin orbitals.

Turning over to CCSD equations, we obtain

$$\Psi_{\text{CCSD}} = e^T \Phi_0$$

where

$$\begin{aligned} T &= T_1 + T_2 \\ T_1 &= \sum_{i,j} t_i^j a_i^\dagger a_j \\ T_2 &= \sum_{i,j,k,l} t_{ik}^{jl} a_i^\dagger a_j a_k^\dagger a_l \end{aligned}$$

Note that  $T_1$  is single excitation operator denoting all configurations where any one electron from the ground state is excited to any of the remaining excited orbitals. The term  $a_i^\dagger a_j$  is removing the electron from the j'th orbital and adding an electron to the i'th orbital. Here we have assumed that j is one of the ground state orbitals and i is one of the higher excited orbitals, unoccupied in the Hartree Fock State. So the action of  $a_i^\dagger a_j$  on the Hartree Fock State "replaces the jth occupied orbital with the i'th unoccupied orbital" (1)

## 3 Quantum Algorithms

### 3.1 Quantum Phase Estimation

#### 3.1.1 Procedure

### 3.2 Variational Quantum Eigensolver

#### 3.2.1 Variational Method

#### 3.2.2 Procedure

## 4 Circuit Construction

### 4.1 Quantum Encoding Methods

How can those second quantized equations be represented in terms of qubits and gates? To bridge classical chemistry and quantum computing, we will first turn to the 3 quantum encoding methods: Jordan Wigner, Parity and Bravyi Kiteav.

#### 4.1.1 Jordan Wigner

In this representation, every qubit corresponds to a spin orbital (SO). The state  $|0\rangle$  represents the SO being unoccupied and the state  $|1\rangle$  represents the SO being occupied. We can also state that the occupation number of an unoccupied SO is 0 and that of an occupied SO is 1. \*\*Possible illustration Insert Here\*\*



How does Jordan Wigner representation help us simulate the electronic hamiltonian? To see those connections, let us convert the qubit states earlier to their matrix representations.

$$\begin{aligned}|0\rangle &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ |1\rangle &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}\end{aligned}$$

What would the creation operator look like if it converts  $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$  to  $\begin{bmatrix} 0 \\ 1 \end{bmatrix}$ ? Using linear algebra and intuition,

$$a^\dagger = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \frac{\sigma^x - i\sigma^y}{2}$$

Similarly,

$$a = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \frac{\sigma^x + i\sigma^y}{2}$$

While we can represent creation and annihilation operators as sum of pauli matrices, these pauli matrices do not preserve the anti-commuting properties of  $a^\dagger$  and  $a$ , namely the fermionic exchange symmetry/antisymmetry principle

$$a_j^\dagger a_k^\dagger = -a_k^\dagger a_j^\dagger$$

but

$$\frac{\sigma_j^x - i\sigma_j^y}{2} \frac{\sigma_k^x - i\sigma_k^y}{2} = \frac{\sigma_k^x - i\sigma_k^y}{2} \frac{\sigma_j^x - i\sigma_j^y}{2}$$

where j and k are 2 arbitrary qubits. To maintain this symmetry within the matrix equation, we recognize the following property of pauli matrices:

$$\sigma^z \sigma^x = -\sigma^x \sigma^z$$

$$\sigma^z \sigma^y = -\sigma^y \sigma^z$$

Utilizing this property, we have to change our representations for  $a^\dagger$

$$a_1^\dagger = \frac{\sigma^x - i\sigma^y}{2} \otimes 1 \otimes 1 \otimes \dots \otimes 1$$

$$a_2^\dagger = \sigma^z \otimes \frac{\sigma^x - i\sigma^y}{2} \otimes 1 \otimes \dots \otimes 1$$

$$a_3^\dagger = \sigma^z \otimes \sigma^z \otimes \frac{\sigma^x - i\sigma^y}{2} \otimes \dots \otimes 1$$

⋮

$$a_n^\dagger = \sigma^z \otimes \sigma^z \otimes \sigma^z \otimes \dots \otimes \frac{\sigma^x - i\sigma^y}{2}$$

(13) At this point, one should choose arbitrary j and k to verify that this matrix representation satisfies the anticommuting property.

There are 2 takeaways from this discussion:

1. Creation and Annihilation Operators can be represented as sum of pauli matrices. Since the hamiltonian is composed of these operators and pauli matrices are quantum gates acting on qubits, we should be comfortable with representation of hamiltonian in terms of matrices or gates acting on qubits.
2. To act on the kth qubit with  $a^\dagger$ , we would need to first apply k-1  $\sigma_z$  gates and then apply  $\frac{\sigma^x - i\sigma^y}{2}$ . Using Big Oh (worst case) Analysis , any electronic operation thus requires  $O(n)$  or linear time.

Can we do better?

#### 4.1.2 Parity

Before we move onto parity basis, it is important to clarify some new terminology. The term “local” in this context means that information that can be extracted from a single qubit alone. On the other hand, the term “nonlocal” means that we need to look at more than one qubit to extract information.

In the Jordan Wigner basis, the occupation number is a local quantity as we only need to look at the kth qubit alone to find its occupation number. However, the parity of kth qubit or the sum of occupation numbers of the first k-1 qubits modulo 2 requires us to check all those k-1 qubits. Thus, parity is a nonlocal quantity in Jordan Wigner Basis.

Now, what is the parity basis? It is best explained using the Jordan wigner basis. Suppose the occupation numbers  $\{o_i\}$  for 4 qubits are the following

$$\begin{bmatrix} o_0 \\ o_1 \\ o_2 \\ o_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 1 \end{bmatrix}$$

Converting from Jordan Wigner basis to Parity basis,

$$\begin{bmatrix} o_0 \mod 2 \\ o_0 + o_1 \mod 2 \\ o_0 + o_1 + o_2 \mod 2 \\ o_0 + o_1 + o_2 + o_3 \mod 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

Thus the qubit  $q_k$  stores the sum of occupation numbers  $o_1 + o_2 + \dots + o_k$  modulo 2. In other words,  $q_k$  stores the parity of the first k-1 qubits. However, if we have a system of n qubits and we change the occupation number  $o_k$  of the  $q_k$ , then such change affects parity of the last  $n - k + 1$  qubits. This is true because each of the qubits in the set  $\{q_k, q_{k+1}, \dots, q_n\}$  store  $o_k$ . Thus, to act on the kth qubit with  $a^\dagger$  involves at least n-k+1 operations. Using Big Oh Analysis, any electronic operation within Parity Basis takes  $O(n)$  or linear time. This is on par with the Jordan Wigner basis. Can we do better?

### 4.1.3 Bravyi Kitaev

Bravyi Kitaev is a combination of keeping track of occupation numbers and parity information both non-locally. This combination of Jordan Wigner and Parity Basis produces a logarithmic running time for action of fermionic operators, which is an improvement over the linear time of previously discussed methods (15).

In this mapping, the even indexed qubits store only their occupation numbers alone, as in Jordan Wigner basis. However, the odd indexed qubits store parity information of a certain set of orbitals. If  $\log i + 1$  is an integer where  $i$  is an odd index, then the  $i$ 'th qubit stores the occupation numbers of all orbitals with indices less than or equal to  $i$  (i.e. sum modulo 2). This is similar to Parity Basis mapping. However, if, for odd index  $i$ ,  $\log i + 1$  is not an integer, then the  $i$ 'th qubit stores occupation numbers of all qubits in a smaller binary set (14). For clarity, we illustrate this with an example of 8 qubit system.

Suppose you have 8 orbitals whose occupation numbers are  $o_1, o_2, \dots, o_8$ . Applying the Bravyi Kitaev transformation, we will obtain the states of all 8 qubits as shown below:

$$\begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} o_0 \\ o_1 \\ o_2 \\ o_3 \\ o_4 \\ o_5 \\ o_6 \\ o_7 \end{bmatrix} = \begin{bmatrix} \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \\ \phantom{0} \end{bmatrix} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{bmatrix}$$

Let's fill in the missing entries of the empty matrix. From earlier discussion, we know that even indexed entries will be assigned their own occupation numbers. Hence,

$$\begin{bmatrix} o_0 \\ o_2 \\ o_4 \\ o_6 \end{bmatrix} = \begin{bmatrix} q_0 \\ q_1 \\ q_3 \\ q_5 \\ q_6 \\ q_7 \end{bmatrix}$$

If the odd index  $i$  makes  $\log(i + 1)$  an integer, it stores occupation numbers of all orbitals with indices less than or equal to  $i$ . In our 8 qubit example, those

qubits would be  $q_1, q_3$ , and  $q_7$ . Thus,

$$\begin{bmatrix} o_0 \\ o_0 + o_1 \\ o_2 \\ o_0 + o_1 + o_2 + o_3 \\ o_4 \\ o_6 \\ o_0 + o_1 + o_2 + o_3 + o_4 + o_5 + o_6 + o_7 \end{bmatrix} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{bmatrix}$$

This leaves us with one missing entry,  $q_5$ . To subdivide into smaller binary set, we note the last odd indexed qubit  $i$  where  $\log(i+1)$  was an integer. In our case,  $i = 3$ . To construct a smaller binary set, we ignore all orbitals with indices  $i \leq 3$ . So our smaller binary set only considers occupation numbers of orbitals indexed 4 and 5. Thus, filling in the missing entry, we obtain

$$\begin{bmatrix} o_0 \\ o_0 + o_1 \\ o_2 \\ o_0 + o_1 + o_2 + o_3 \\ o_4 \\ o_4 + o_5 \\ o_6 \\ o_0 + o_1 + o_2 + o_3 + o_4 + o_5 + o_6 + o_7 \end{bmatrix} = \begin{bmatrix} q_0 \\ q_1 \\ q_2 \\ q_3 \\ q_4 \\ q_5 \\ q_6 \\ q_7 \end{bmatrix}$$

One can easily observe that each occupation number  $o_i$  is present in at most  $\log(n)+1$  entries. Thus, the action of any fermionic operator need only  $O(\log(n))$  qubit operations, which is better than the other methods.

#### 4.1.4 Application

Why Jordan Wigner/Parity instead of Bravyi Kitaev? Use Jordan Wigner to convert Hamiltonian into pauli notation

## 4.2 Hamiltonian Decomposition

### 4.2.1 Trotterization

### 4.2.2 Exponentiation

## 5 Variational Quantum Eigensolver

The goal of this algorithm is to minimize the ground state energy of the system by choosing a certain set of cluster amplitudes  $t$ .

## 5.1 State Preparation

In this part, we prepare first the reference wavefunction and then we apply the  $U(t)$  unitary to prepare the UCCSD wavefunction or the UCCSD variational form (4).

### 5.1.1 Reference Wavefunction

For systems with weak correlations, Hartree Fock state provides a good reference wavefunction because its easy to prepare and has a high overlap with the exact ground state. Such overlap would make convergence to the exact ground state easier for the algorithm (4)(7). The ease of preparing the Hartree Fock state can be seen from the following equation:

$$\Phi_0 = a_n^\dagger a_{n-1}^\dagger \cdots a_1^\dagger |0\rangle$$

where  $\Phi_0$  is the Hartree fock state of an n electron system. This is equivalent to filling up the ground state orbitals and leaving the other N-n orbitals unoccupied. If we were to use Jordan Wigner mapping, this would  $\Phi_0$  would correpond to the state  $|0\rangle^{\otimes N-n} \otimes |1\rangle^{\otimes n}$  (4).

Hole: What about strongly correlated systems?

### 5.1.2 UCCSD Wavefunction

Once the reference state has been prepared, we then apply  $U(t)$  to that state to obtain the UCCSD wavefunction. Recall from earlier section that

$$U(t) = e^{T-T^\dagger}$$

Since  $U(t)$  can be represented by string of pauli matrices (4), we only need to apply a few quantum gates to qubits after preparing reference state.

Where do the initial param come from? 2nd order Moller Plesset Theory  
\*\*\*Thorough explanation needed HERE\*\*\*

## 5.2 Energy Measurement

Energy Measurement can be done in 2 ways: Quantum Phase Estimation and Hamiltonian Averaging. Let's tackle QPE first. Since we have a state prepared and the hamiltonian operator, we can use QPE to measure the phase of the eigenvalue of the hamiltonian and thus obtain the energy approximation. There are 2 shortcomings of this approach. First, QPE requires long coherence times (the time before a qubit becomes susceptible to error) which are impractical with current quantum computing technology. Second, if the state is a combination of eigenstates, then we would have to repeat the measurement process a bunch of times which is also not practical !!(10)

A good alternative is the Hamiltonian Averaging Procedure. Recall the qubit hamiltonian from the basis set discussion as a sum of weighted pauli matrices.

$$H = \sum_{i_1 \alpha_1} h_{\alpha_1}^{i_1} \sigma_{\alpha_1}^{i_1} + \cdots$$

Using Linearity of Expectation,

$$\langle H \rangle = \sum_{i_1 \alpha_1} h_{\alpha_1}^{i_1} \langle \sigma \rangle_{\alpha_1}^{i_1} + \dots$$

Hence measuring the expectation value of the hamiltonian reduces to measuring the expectation value of the individual pauli terms. This measurement of expectation value of each term  $O_i$  will require  $n_i$  iterations of state preparation and measurement. Here, the number of iterations  $n_i$  is dependent upon the variance of  $\langle O_i \rangle$ , the weight of that pauli term  $h_i$  and the desired precision  $\epsilon$  (10). Once we have measured the expectation values of all the pauli terms, we can scale the terms based on their weight  $h_i$  and sum them to obtain the expectation value of the qubit hamiltonian. Hence, through hamiltonian averaging, we can obtain the energy of the state corresponding to a given set of parameters.

How do we measure each individual term? Let's say we have a 4 qubit system and the term

$$O = \sigma_3^x \sigma_2^z \sigma_1^y \sigma_0^z$$

Since the eigenstates of  $\sigma^z$  are the computational basis states  $|0\rangle$  and  $|1\rangle$ , we just apply the measurement gate to the qubits  $|q_2\rangle$  and  $|q_0\rangle$ . However, for the other qubits  $|q_3\rangle$  and  $|q_1\rangle$ , a measurement gate is not enough. For  $|q_3\rangle$ , we have to first apply a hadamard gate to convert from computational basis  $|0\rangle$  and  $|1\rangle$  to eigenvectors of  $\sigma^x$  which are  $(|0\rangle + |1\rangle)/\sqrt{2}$  (bell states). Similarly, for  $|q_1\rangle$ , we apply the rotation operator  $R_x(\pi/2)$  to convert to the eigenstates of  $\sigma^y$  which are  $(|0\rangle + i|1\rangle)/\sqrt{2}$  and  $(i|0\rangle + |1\rangle)/\sqrt{2}$ . These additional operations project the computational basis onto the eigenbasis of  $\sigma^i$  and are thus called projective measurements. The final circuit is shown below.

\*\*\* Insert Figure \*\*\*

### 5.3 Parameter Optimization

Once energy  $E(\vec{\theta}_1)$  is measured, we need to find a new set of guesses  $\vec{\theta}_2$  for the unitary coupled cluster amplitudes  $\vec{t}$  to minimize the energy. Since equations to find  $\vec{t}$  are nonlinear (as discussed in Coupled Cluster section), we have to need a nonlinear optimization algorithm (4). There are 2 classes of such algorithms: Direct Search and Gradient Bases Methods.

#### 5.3.1 Direct Search

In these methods, we start with an initial guess  $\vec{\theta}_1$  as starting point and evaluate the energy at all points around that starting point (16). The algorithm then chooses the point  $\vec{\theta}_2$  which minimizes the energy the most relative to the energy of the starting point  $E(\vec{\theta}_1)$ , and  $\vec{\theta}_2$  becomes starting point in the next iteration. Eventually, the algorithm finds the stationary point  $E(\vec{\theta}_n)$  that minimizes energy.

Being less susceptible to quantum noise, Direct Search Algorithms were preferred over Gradient Based Methods. Indeed, quantum noise makes the objective function or electronic energy (as a function over parameters  $\vec{\theta}$ ) discontinuous, making Direct Search Algorithms even more favorable. However, these algorithms require too many function evaluations or too many quantum measurements, resulting in their inefficiency as quantum computers improve (4).

### 5.3.2 Gradient Based Methods

## 5.4 Qiskit Implementation

# 6 Quantum Phase Estimation

## 6.1 Inverse Fourier Transform

## 6.2 Phase Analysis

# 7 References

1. <https://ebookcentral-proquest-com.proxy.libraries.rutgers.edu/lib/rutgers-ebooks/reader.action?docID=4751479>
2. Cite wikipedia on electron correlation on “instant”
3. <https://chemistlibrary.files.wordpress.com/2015/02/modern-quantum-chemistry.pdf>
4. McClean’s paper (4) (strategies for quantum computing molecular energies)
5. <https://qiskit.org/textbook/ch-applications/vqe-molecules.htmlbackgroundmath>
6. <https://www.cs.umd.edu/class/fall2019/cmsc657/projects/group12.pdf>*null*
7. *null*
8. *null*
9. *null*
10. Theory of VQE paper <https://arxiv.org/abs/1509.04279>
11. Cite (11) Wikipedia on Multireference configuration interaction
12. <https://arxiv.org/pdf/1304.3061.pdf> Photonic Processor
13. <https://docs.microsoft.com/en-us/quantum/libraries/chemistry/concepts/jordan-wigner?view=qsharp-previewf>
14. <https://arxiv.org/pdf/1812.02233.pdf> More recent version
15. <https://arxiv.org/abs/1208.5986> Seeley