Alma Mater Studiorum · University of Bologna

School of Science
Department of Physics and Astronomy
Master Degree in Physics

Excited states calculations with the Quantum Approximate Optimization Algorithm

Supervisor:
Prof. Elisa Ercolessi

Submitted by:
Daniele Trisciani

Co-supervisor:
Dr. Claudio Massimiliano Sanavio

Academic Year 2021/2022

The saddest aspect of life right now is that science gathers knowledge faster than society gathers wisdom.

Isaac Asimov

Abstract

Il Quantum Approximate Optimization Algorithm (QAOA) è un algoritmo ibrido classico-quantistico, introdotto per risolvere problemi di ottimizzazione combinatoria, come il $Max\ Cut$. Il QAOA sfrutta un circuito quantistico parametrico per stimare lo stato fondamentale di una Hamiltoniana costo H_c, che codifica la soluzione del problema combinatorio. I migliori valori per i parametri sono determinati tramite tecniche di ottimizzazione classica. Il QAOA è anche usato per ottenere lo stato fondamentale delle molecole. In questo lavoro, noi estendiamo la sua applicabilità introducendo una procedura che ci permette di calcolare anche gli stati eccitati in modo iterativo: una volta che lo stato fondamentale è conosciuto, è possibile ottenere una nuova Hamiltoniana cui stato fondamentale è il primo stato eccitato dell'Hamiltoniana originale. Una volta fatto il QAOA può essere applicato di nuovo. Il metodo proposto è stato testato sulle molecole di H_2 e LiH. Per la parte quantistica dell'algoritmo, la base STO-3G è stata usata per il calcolo degli integrali a un corpo e a due corpi, considerando solo 2 orbitali molecolari e 2 elettroni con spin opposto, un step necessario per ottenere l'Hamiltoniana di seconda quantizzazione. Gli operatori di creazione e distruzione sono stati mappati sui qubit usando le trasformazioni di Jordan Wigner. Riguardo all'ottimizzatore classico, sono stati usati il metodo Basin-Hopping con l'algoritmo BFGS. E' prima stata calcolata l'energia e la funzione d'onda dello stato fondamentale. Per entrambe le molecule, gli stati fondamentali possono essere stimati con successo per piccole distance inter-nucleari. A grandi valori i risultati non sono buoni, ciò è dovuto alla degenerazione tra lo stato fondamentale e il primo stato eccitato. Poi la procedura è applicata per il primo stato eccitato per entrambe le molecole. Come prima, otteniamo un ottima convergenza solamente a corte distanze inter-nucleari. Il problema della degenerazione non è più presente, ma gli errori legati al calcolo dello stato fondamentale sono strasferiti al calcolo del primo stato eccitato.

Abstract

The Quantum Approximate Optimization Algorithm (QAOA) is a hybrid quantum-classical algorithm introduced to solve complex combinatorial optimization problems, such as the $Max\ Cut$. It exploits a parameterized quantum circuit to estimate the ground state of a cost Hamiltonian H_c, that encodes the solution of the combinatorial problem. The best values of the parameters are determined via classical optimization techniques. The QAOA is also used to obtain the ground state of molecules. In this work, we will extend its applicability by introducing a procedure that allows us to calculate also the excited states in an iterative way: once the ground state is known, one can obtain a new Hamiltonian whose ground state is the first excited state of the original Hamiltonian. Now the QAOA can be applied again. The proposed method has been tested using the H_2 and the LiH molecule. For the quantum part of the algorithm, the STO-3G basis has been used for the one-body and two-body integral calculation, considering only 2 molecular orbitals and 2 electrons with opposite spins, a necessary step to build the second quantized Hamiltonian. The creation and annihilation operators have been mapped to qubits using the Jordan Wigner transformation. For the classical optimizer, the Basin-Hopping method with the BFGS algorithm has been used. We first calculate the ground state energy and wave function. For both the molecules, the ground states can be successfully estimated for small inter-nuclear distances. At larger values the results are not good, due to the degeneration between the ground state and the first excited state. Then we apply our procedure for the calculation of first excited state for both molecules. Again, we show convergence to the right results for short inter-nuclear distances. The degeneration problem is no longer present, but the errors related to the calculation of the ground states are transferred to the calculation of the first excited states.

Contents

Introduction

Classical computers are now part of our lives and have drastically changed society by having the ability to solve otherwise unsolvable problems. Unfortunately some problems are impractical even for supercomputers. A quantum computer is a computer that uses the laws of quantum mechanics to solve this kind of problems. In contrast to the bit, the basic unit of information is called qubit, which has the peculiarity to exist in a superposition of states. In the quantum circuit model, the qubit is manipulated using quantum logic gates. A quantum algorithm defines a quantum circuit, composed by a sequence of quantum gates, that acts on some input qubits and terminates with a measurement. Relevant quantum algorithms are the Shor's algorithm, which promises to perform integer factorization in a feasible amount of time and the Grover algorithm, developed to speed up the unstructured search. Quantum computation finds application in many problems such as computational chemistry, drug design, cryptography, machine learning.

Quantum supremacy, the goal of proving that a programmable quantum device can solve a problem that no classical computer can solve in a reasonable amount of time, has been achieved in 2019 by Google Quantum AI using 54 qubits. Unfortunately it is estimated that a lot of quantum algorithms require more the 1000 qubits to outperform their classical counterpart. Indeed quantum computers suffer from decoherence, when a quantum system loses its quantum properties, and quantum noise, which originates from various sources, i.e., the environment. Quantum Error Correction (QEC) has been developed for this reason. It is composed of a set of tools for the correction of quantum errors and is essential to achieve quantum supremacy. But error correction methods require additional qubits, and nowadays the number of qubits that can be implemented is limited.

Existing quantum computers are called Noisy Intermediate-Scale Quantum (NISQ) devices. They are composed of about 100 noisy qubits. Variational quantum algorithms have been developed for this kind of devices. They are hybrid quantum-classical algorithms, which delegate the classically difficult part of a problem to the quantum computer and perform the others on some classical device. In particular, these algorithms use a parameterized quantum circuit (PQC) $U(\vec{\theta})$ that depends on some parameters $\vec{\theta}$, usually randomly initialized. The PQC is applied to the qubit states that are used for the

measurement of some cost Hamiltonian H_c. The result of the measure is used for the optimization of the parameters $\vec{\theta}$ by a classical device, using an optimization method such as the stochastic gradient descent, therefore a new PQC $U(\vec{\theta}_{new})$ is obtained. The whole procedure is iterated until convergence is established.

The most famous VQA algorithms are the Variational Quantum Eigensolver (VQE), which has found applications in computational chemistry and quantum machine learning, and the Quantum Approximate Optimization Algorithm (QAOA), which has been used in this work.

The QAOA is a quantum algorithm developed to solve combinatorial problems such as the $Max\ Cut$. It uses a PQC characterized by a unitary operator $U(\vec{\beta}, \vec{\gamma})$ which is given in function of a set of parameters $(\vec{\beta},\ \vec{\gamma})$, whose number is proportional to an integer $p \geq 1$. Increasing the value of p leads to better results, but the depth of the circuit is also increased. The PQC is composed by the alternate application of two unitaries $exp\{-i\gamma_i H_c\}$ and $exp\{-i\beta_i H_m\}$, with $1 \leq i \leq p$, where H_c and H_m are respectively called the cost and mixing Hamiltonian. H_c is chosen so that the solution of the problem is represented by its ground state, while H_m is chosen to satisfy $[H_m, H_c] \neq 0$, so that the total evolution is able to to span the whole Hilbert space. At the beginning of the algorithm, the parameters $(\vec{\beta}_0,\ \vec{\gamma}_0)$ are randomly initialized and the initial states is fixed to an equal superposition of the computational basis states. Then, the PQC evolves the latter to $|\psi(\vec{\beta}_0, \vec{\gamma}_0)\rangle$ which is used for the estimation of the cost function $\langle\psi(\vec{\beta}_0, \vec{\gamma}_0)| \, |H_c| \, |\psi(\vec{\beta}_0, \vec{\gamma}_0)\rangle$. The classical optimizer exploits this measure to find new parameters $(\vec{\beta}_1,\ \vec{\gamma}_1)$ for which the ground state is better estimated. This procedure is repeated until the convergence of the optimizer is reached. As result we obtain the ground state $|\psi_{g.s.}\rangle = |\psi(\vec{\beta}_{opt}, \vec{\gamma}_{opt})\rangle$ and the ground state energy $\alpha_{g.s.} = \langle\psi_{g.s.}| \, H_c \, |\psi_{g.s.}\rangle$ of the cost Hamiltonian.

The QAOA algorithm has been used only to estimate the ground state of a Hamiltonian, this is very limiting if the excited states are of interest. This work illustrates a procedure for calculating the excited states using the QAOA algorithm.

The procedure is tested using the H_2 molecule and the LiH molecule. The STO-3G basis set has been used to calculate the one-body and two-body integral, considering only 2 molecular orbitals and 2 electrons with opposite spins, a necessary step for constructing the second quantized Hamiltonian. The creation and annihilation operators, $a^\dagger$ and a, have been mapped using the Jordan-Wigner transformation. The Basin-Hopping method with the BFGS algorithm has been used for the classical optimization. As a prerequisite, the ground states of the two molecules are also calculated.

The content of the thesis is as follows.

In Chap. 1 the basic quantum properties for the treatment of quantum computation will be introduced, followed by a description of quantum gates, such as the single qubit gates and the controlled gates. Considering the circuit model, the quantum algorithms are composed of these quantum gates. Finally, we will briefly introduce the requirements

for a physical realization of a quantum computer, and the steps scientists are taking to reach quantum supremacy.

In Chap. 2 we will introduce Variation Quantum Algorithms (VQA). These types of algorithms are designed for NISQ devices, and they are referred to as hybrid quantum-classical algorithms. Then we will focus on one type of VQA algorithm, the Quantum Approximate Optimization Algorithm, followed by the explanation of some optimization algorithms, such as the BFGS and the Basin-Hopping method.

The benchmark of the procedure for calculating excited state using the QAOA will be done using diatomic molecule Hamiltonians. For this reason, in Chap. 3 we will see how to obtain the Schrodinger equation for the H atom and then diatomic molecules. Finally we will introduce approximation for computational chemistry, and what is the role of quantum computation in chemistry nowadays.

In Chap. 4 the procedure for the calculation of excited states will be introduced. We will apply the QAOA to calculate of the ground state of the H_2 and LiH molecules for some distances, a prerequisite of the procedure. After applying a transformation to the Hamiltonian, the first excited state will be calculated using the QAOA, and the results will be benchmarked with respect to the calculations of exact diagonalization.

Chapter 1

Introduction to Quantum Computing

In this chapter, we will introduce the basic concepts of quantum computation, starting with the quantum properties of the qubit, the basic unit of information in quantum computation. Then we will focus on the basic building blocks of a quantum circuit, quantum gates, a universal language for describing algorithms. In the end, we will talk about the properties that an experimental setup must satisfy to be called a quantum computer, and the steps the research is taking towards a fault-tolerant quantum computer. The main reference used for this part is [4].

1.1 Qubits and Hilbert space

Quantum computers exploit quantum mechanical phenomena to manipulate information. The basic unit of information in classical computation is the bit. Similarly to a coin, it can be in two states, 0 (the head) or 1 (the tail). On the other hand, the basic unit of information in quantum computation is the qubit.

A qubit is described as a quantum mechanical object, which, in analogy to a classical bit, may be in the states $|0\rangle$ and $|1\rangle$. But, unlike the bit, the state $|\psi\rangle$ of the qubit may also be a linear superposition, of the states $|0\rangle$ and $|1\rangle$:

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle \tag{1.1}$$

where the coefficients $\alpha, \beta \in \mathbb{C}$, and $|\alpha|^2 + |\beta|^2 = 1$. The state of a qubit is a unit vector in a two-dimensional complex Hilbert space $\mathscr{H}$. The special states $|0\rangle$ and $|1\rangle$ form an orthonormal basis known as *computational basis*.
When a classical computer retrieves data from its memory, the value of a bit is always 0 or 1. On the other hand, when we try to measure the value of a qubit in the state $|\psi\rangle$,

we obtain either 0 with probability $|\alpha|^2$ or 1 with probability $|\beta|^2$. It turns out that the only way to know α and β is to measure infinitely identically prepared qubits.

The geometric interpretation of a qubit can be better grasped if the equation (1.1) is rewritten, up to a global phase, as

$$|\psi\rangle = \cos\frac{\theta}{2}|0\rangle + e^{i\varphi}\sin\frac{\theta}{2}|1\rangle, \tag{1.2}$$

where θ and φ describe the position on a unitary sphere. Indeed we can visualize the possible quantum states on the Block sphere, shown in Figure 1.1.

The evolution of a closed quantum system is described by a unitary transformation. Given the state $|\psi\rangle$ at time t_1, it evolves to the state $|\psi'\rangle$ at time t_2 through a unitary operator $U(t_1, t_2)$,

$$|\psi'\rangle = U(t_1, t_2)|\psi\rangle. \tag{1.3}$$

Applying an unitary operator on a quantum state changes it, but an observer cannot get any information about the states. To achieve that, a quantum measurement must be performed.

Measurements

Quantum measurements are described by a collection $\{M_m\}$ of measurements operators, where the probability that the result m occurs is given by

$$p(m) = \langle\psi\,|M_m^\dagger M_m|\,\psi'\rangle \tag{1.4}$$

hence the state of the wave function collapse into the new state

$$|\psi\rangle \to \frac{M_m|\psi\rangle}{\sqrt{p(m)}} \tag{1.5}$$

The property of the probabilities to sum to one is assured by the *completeness equation*

$$\sum_m M_m^\dagger M_m = I \tag{1.6}$$

The measurement of a qubit in the computational basis is particularly relevant, its defined by the two measurements operators $M_0 = P_0 = |0\rangle\langle 0|$, $M_1 = P_1 =|\,1\rangle\langle 1|$, the projectors on the computational basis states $|0\rangle$ and $|1\rangle$. These measurements are Hermitian and $M_0^2 = M_0$, $M_1^2 = M_1$. Thus, when a qubit is measured in the computational basis, that measurements leads to outcome an 0, corresponding to the state $|0\rangle$, with probability $|\alpha|^2$ and an outcome 1, corresponding to the state $|1\rangle$, with probability $|\beta|^2$.

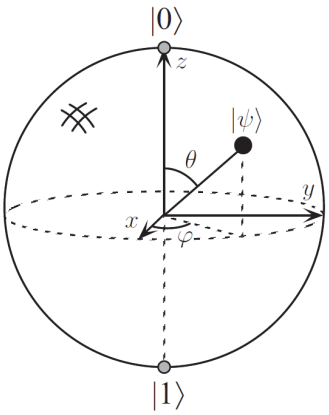

Figure 1.1: Bloch sphere representation of a qubit. From [4]

A projective measurement is described by an observable, M, with spectral decomposition,

$$M = \sum_m m P_m \tag{1.7}$$

where P_m is the projector onto the eigenspace of M with eigenvalue m. Upon measuring the state $|\psi\rangle$, the probability of getting result m is given by

$$p(m) = \langle \psi | P_m | \psi \rangle . \tag{1.8}$$

Given that outcome m occured, the state of the quantum system immediately after the measurement is

$$\frac{P_m|\psi\rangle}{\sqrt{p(m)}} \tag{1.9}$$

When one moves from theory to experiment, the interaction between these states and the outside world must be considered. It brings random interactions called random noise that cannot be completely neglected. The control of such noise processes are the only way to build useful quantum information processing systems.

An operation that will be useful later is the trace. Given a matrix A, the trace of A is defined to be the sum of its diagonal elements

$$\mathrm{tr}(A) \equiv \sum_i A_{ii}. \tag{1.10}$$

it has properties such as linearity and cyclicity:

$$\mathrm{tr}(x\mathbf{A} + y\mathbf{B}) = x\,\mathrm{tr}(\mathbf{A}) + y\,\mathrm{tr}(\mathbf{B}), \quad \mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA}) \tag{1.11}$$

In general many qubits are required for computational tasks, so we have to consider multiple qubit systems. The state space of a quantum system made up of more than one distinct physical system is called *composite system*.

Density operator

Another possible formulation for the description of a quantum system, alternatively to the use of state vectors, is the use of the density operator also called density matrix. This formulation, although is mathematically equivalent to the previous one, its is able to describe a quantum system as a mixed state, therefore, is able to take in account the presence of noise. A mixed state can be described by the ensemble of pure states $\{p_i, |\psi_i\rangle\}$, where p_i is the probability of the state $|\psi_i\rangle$. The density operator for the system is defined by

$$\rho \equiv \sum_i p_i |\psi_i\rangle \langle\psi_i|. \tag{1.12}$$

Unitary evolution and measurements are easily described in the language of density operators; for proof see [4]. Thus, given a unitary operator U, the evolution of the density operator is described by

$$\rho' = \sum_i p_i U |\psi_i\rangle \langle\psi_i| U^\dagger = U\rho U^\dagger \tag{1.13}$$

and, for a measurement described by measurement operators M_m, the probability of obtaining result m is

$$p(m) = \mathrm{tr}\left(M_m^\dagger M_m \rho\right) \tag{1.14}$$

the state after the measurement which yields the result m described by the density operator is equal to

$$\rho_m = \frac{M_m \rho M_m^\dagger}{p(m)}. \tag{1.15}$$

The state space $\mathscr{H}$ of a composite physical system is the tensor product of the state spaces of the component physical systems. Thus, if we have systems numbered from 1 to n, and system number i is prepared in the state ρ_i, then the joint state of the total system is $\rho_1 \otimes \rho_2 \otimes \ldots \otimes \rho_n$. In general, however, a state $p \in \mathscr{H}$ cannot be written as a tensor product of states of individual systems.

The density operator ρ associated to some ensemble $\{p_i, |\psi_i\rangle\}$ must satisfies the conditions:

1. (**Positivity condition**) ρ is a positive operator.

2. (**Trace condition**) $\mathrm{tr}(\rho) = 1$.

For prof see [4].

A density operator can describe a subsystem of a composite quantum system via the *reduced density operator*. In a system, composed by the subsystems A and B, described by ρ^{AB}, the reduced density operator for system A is defined by

$$\rho^A \equiv \mathrm{tr}_B\left(\rho^{AB}\right) \tag{1.16}$$

where tr_B is known as *partial trace* over the system B, with the definition

$$\mathrm{tr}_B(|a_1\rangle_A \langle a_2|_A \otimes |b_1\rangle_B \langle b_2|_B) \equiv |a_1\rangle_A \langle a_2|_A \,\mathrm{tr}(|b_1\rangle_B \langle b_2|_B) \tag{1.17}$$

where $\{|a_i\rangle_A\}$ and $\{|b_j\rangle_B\}$ form an orthonormal basis in $\mathscr{H}_A$ and $\mathscr{H}_B$ respectively.

Fidelity

A tool for the measure of distance between quantum states is the fidelity. The fidelity is defined, given any two states ρ and σ as,

$$F(\rho, \sigma) \equiv \mathrm{tr}\sqrt{\sqrt{\rho}\sigma\sqrt{\rho}}$$

With the properties:

- **Symmetry** $F(\rho, \sigma) = F(\sigma, \rho)$
- **Bounded values** $0 \leq F(\rho, \sigma) \leq 1$, and $F(\rho, \rho) = 1$.

At first glance, it doesn't seem a useful measure of the distance between the states ρ and σ, and also, it doesn't seem symmetric. However, we can convince ourselves, considering the case when ρ and σ commute, i.e., they are diagonal in the same basis:

$$\rho = \sum_i t_i |i\rangle\langle i|; \quad \sigma = \sum_i w_i |i\rangle\langle i| \tag{1.18}$$

In this case the fidelity takes the form:

$$F(\rho, \sigma) = \mathrm{tr}\sqrt{\sum_i t_i w_i |i\rangle\langle i|} = \mathrm{tr}\left(\sum_i \sqrt{t_i w_i}|i\rangle\langle i|\right) = \sum_i \sqrt{t_i w_i} = F\left(t_i, w_i\right). \tag{1.19}$$

Now the fidelity is explicitly symmetric, and depends only on the eigenvalues.

1.2 Quantum gates and quantum circuits

Classical computer circuits consist of wires used to carry information around and logic gates that manipulate the information according to an algorithm. Similarly, a quantum computer is built from a *quantum circuit* that form a collection of *quantum gates*, while the latter carries around and manipulates the quantum information.

Single qubit gates

Recalling that $|\alpha|^2 + |\beta|^2 = 1$, operations on a qubit must be unitary transformations in order to preserve its unitary norm. Some of the most important ones are the Pauli matrices:

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

The X Pauli matrix defines the quantum NOT gate, that interchanges the role of $|0\rangle$ and $|1\rangle$ of a state

$$X \begin{bmatrix} \alpha \\ \beta \end{bmatrix} = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}. \tag{1.20}$$

The Z Pauli gate leaves the direction $|0\rangle$ unchanged, but flips the sign of the qubit in direction $|1\rangle \to -|1\rangle$. Instead the application of the Y Pauli gate performs a rotation of π radians around the y-axis.

Three other fundamental single qubit gates are the Hadamard gate H, the phase gate S and the phase gate T (historically known as $\pi/8$ gate):

$$H \equiv \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}; \quad S \equiv \begin{bmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{2}} \end{bmatrix}; \quad T \equiv \begin{bmatrix} 1 & 0 \\ 0 & e^{i\frac{\pi}{4}} \end{bmatrix}.$$

The Hadamard gate rotates the sphere around the $\hat{y}$ axis by 90°, followed by a rotation around the $\hat{x}$ axis by 180°. The phase gates S and T act on the qubits such that $|0\rangle \mapsto |0\rangle$ and $|1\rangle \mapsto e^{i\varphi}|1\rangle$, where $\varphi = \pi/2$ or $\pi/4$ respectively, leaving unchanged the probability to measure this states. Symbols and definitions for these single qubit gates are shown in Figure 1.2.

The Pauli matrices are the generators of three transformations, the rotation operators about the $\hat{x}$, $\hat{y}$ and $\hat{z}$ axes of the Bloch sphere, and they are defined by:

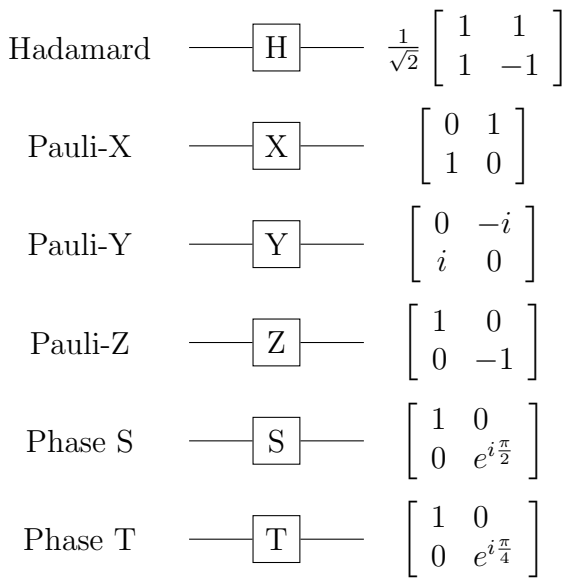

Figure 1.2: Symbols for the common single qubit gates.

$$R_x(\theta) \equiv e^{-i\theta X/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}X = \begin{bmatrix} \cos\frac{\theta}{2} & -i\sin\frac{\theta}{2} \\ -i\sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \tag{1.21}$$

$$R_y(\theta) \equiv e^{-i\theta Y/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}Y = \begin{bmatrix} \cos\frac{\theta}{2} & -\sin\frac{\theta}{2} \\ \sin\frac{\theta}{2} & \cos\frac{\theta}{2} \end{bmatrix} \tag{1.22}$$

$$R_z(\theta) \equiv e^{-i\theta Z/2} = \cos\frac{\theta}{2}I - i\sin\frac{\theta}{2}Z = \begin{bmatrix} e^{-i\theta/2} & 0 \\ 0 & e^{i\theta/2} \end{bmatrix} \tag{1.23}$$

For any real numbers α, β, γ and δ we have the following relation:

$$U = e^{i\alpha}R_z(\beta)R_y(\gamma)R_z(\delta) \tag{1.24}$$

where U is a unitary operation on a single qubit. For prof see [4]

Hadamard basis

The states $|0\rangle$ and $|1\rangle$ are not the only possible choice of basis states. Another basis states very popular in quantum information theory, called Hadamard basis, is the set of states

$$|+\rangle \equiv \frac{(|0\rangle + |1\rangle)}{\sqrt{(2)}} \quad \text{and} \quad |-\rangle \equiv \frac{(|0\rangle - |1\rangle)}{\sqrt{2}}.$$

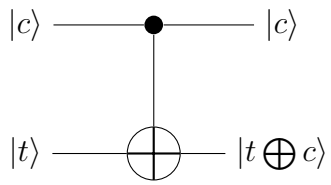

Figure 1.3: Circuit representation for the controlled-NOT gate. The top line represents the control qubit, the bottom line the target qubit.

If a qubit is measured with respect to this new basis, it results in the result "+" with probability $|\alpha+\beta|^2/2$ and the result "-" with probability $|\alpha-\beta|^2/2$, with corresponding post-measurement states $|+\rangle$ and $|-\rangle$, respectively.
An generic state $|\psi\rangle = \alpha\,|0\rangle + \beta\,|1\rangle$ can be written in terms of the Hadamard basis as

$$|\psi\rangle = \alpha|0\rangle + \beta|1\rangle = \alpha\frac{|+\rangle + |-\rangle}{\sqrt{2}} + \beta\frac{|+\rangle - |-\rangle}{\sqrt{2}} = \frac{\alpha+\beta}{\sqrt{2}}|+\rangle + \frac{\alpha-\beta}{\sqrt{2}}|-\rangle. \tag{1.25}$$

Therefore it is possible to treat the Hadamard basis as they were the computational basis states. The Hadamard gate, introduced before, performs the change of basis in a quantum circuit.

Controlled gates

Controlled operations are fundamental in classical computing; the same is valid in quantum computation; one example of a controlled operation is the controlled-NOT (CNOT) gate. This gate is applied to 2 qubits, they are called the control qubit ($|c\rangle$) and the target qubit ($|t\rangle$). The circuit is represented in Figure 1.3. In terms of the computational basis, the CNOT performs the transformation $|c\rangle|t\rangle \to |c\rangle|t \oplus c\rangle$, where $|c\rangle, |t\rangle \in \{|0\rangle, |1\rangle\}$ and $\oplus$ is the XOR operation. Therefore, when $|c\rangle = |0\rangle$, $|t\rangle$ is left unchanged, but if $|c\rangle = |1\rangle$ a bit flip is applied to $|t\rangle$. The CNOT gate, respect to a two qubits basis, is represented by the unitary trasformation

$$U_{CN} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}. \tag{1.26}$$

Generalizing, if U is a single qubit unitary operation, the *controlled-U* operation is a two qubit operation, with a control and a target qubit as before, where, if the control qubit is $|1\rangle$, then U is applied to the target qubit, otherwise, and the target qubit is left unchanged, the transformation is

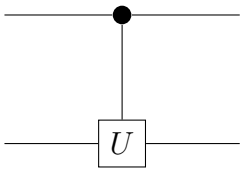

Figure 1.4: Controlled-U operation. The top line is the control qubit, and the bottom line is the target qubit.

$$|c\rangle_1|t\rangle_2 \to \left(\frac{1+Z_1}{2} + \frac{1-Z_1}{2}U_2\right)|c\rangle_1|t\rangle_2. \tag{1.27}$$

The *controlled-U* operation is represented by the circuit shown in Figure 1.4.

Moving to a three-qubit system, the analog of the CNOT gate is the Toffoli gate (also called CCNOT), it has two control qubits instead of one. The Toffoli gate is associated to the unitary trasformation

$$\text{CCNOT} \equiv (II - |11\rangle\langle 11|) \otimes I + |11\rangle\langle 11| \otimes X = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}. \tag{1.28}$$

Now if the two control qubit $|c\rangle_1$ and $|c\rangle_2$ are equal to $|1\rangle$, the target qubit $|t\rangle$ is flipped, if only one control qubit is $|0\rangle$, $|t\rangle$ is left unchanged.

SWAP gate

Another important quantum gate in quantum computing is the SWAP gate, that exchanges the state of two qubits. In the computational basis, the SWAP gate maps

$$|00\rangle \to |00\rangle,\ |01\rangle \to |10\rangle,\ |10\rangle \to |01\rangle \text{ and } |11\rangle \to |11\rangle$$

The SWAP gate is equivalent to the unitary transformation

$$\text{SWAP} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

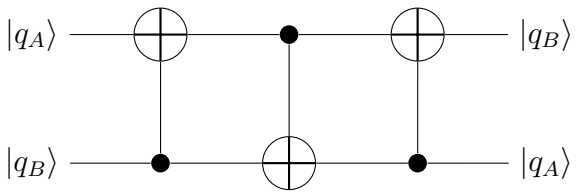

Figure 1.5: Circuit representation for the SWAP gate. The qubit q_A and q_B are swapped.

with respect to the computational basis. It can be decomposed in 3 CNOTS, as showed in figure 1.5.

Universal quantum gates

In classical computation, Boolean functions are expressed using the gates AND, OR, NAND, NOT etc. The NAND gate is called universal because it can compose the NOT, OR and AND gates and from there, anything else. The NAND gate therefore is sad to be Universal. A similar result can be achieved for quantum computation. If a set of quantum gates can approximate any unitary operation with arbitrary accuracy, we can say they are universal for quantum computation. We can find a possible set of gates starting from the assumption that any unitary operators can be exactly expressed as the product of a CNOT and a single qubit gate, the latter can be approximated by the Hadamard, phase and $\pi/8$ gates, for proof see [4]. We can conclude that any unitary operation can be approximated using the Hadamard, phase, $\pi/8$ and CNOT gates. One might ask how many gates are necessary to compose a unitary transformation. Unitary operators exist for which exponentially many gates are required the approximation.

Measures of Pauli strings

The Pauli string representation can be used to efficiently parameterize the observable. The measurement of Pauli string works in this way, the expectation value of σ_z on a qubit is $\langle\sigma_z\rangle = |\alpha|^2 - |\beta|^2$ where $|\alpha|^2$ is the probability for the state $|0\rangle$ and $|\beta|^2$, the Pauli matrices σ_x and σ_y can be transformed into the σ_z basis using the phase gate S and the Hadamard gate H. Therefore Pauli strings $\hat{P}$ on qubits $n \in N$ can be measured using the same procedure on each qubit as

$$\langle\hat{P}\rangle_U = \langle\prod_{k\in K}\sigma_z(k)\rangle_{\tilde{U}U} \tag{1.29}$$

where $\tilde{U}$ is a product of single qubit rotations that depends on the Pauli matrix at qubit k.

1.3 Quantum computers

The physical realization of quantum qubits and gates has been proven to be extremely complicated. To realize a quantum computer, DiVincenzo's criteria [20] must be met:

1. A physical system with well-characterized qubit.

   Where "well characterized" means that the Hamiltonian of the qubit, and all the possible couplings to external fields used to evolve the state of the qubit, have to be known.

2. The capability to prepare the initial state of the qubits to a simple fiducial state.

   Quantum computing requires the initialization of the qubits at a known value before the start of the computation.

3. Decoherence times much longer than the gate operation time.

   Decoherence transforms a quantum system into a classical one; therefore it is essential that the execution of a quantum gate is faster than the decoherence time.

4. A universal set of quantum gates.

   A quantum algorithm can be seen as a sequence of unitary transformations (quantum gates) $U_1 = e^{iH_1 t}$, $U_2 = e^{iH_2 t}, ...$ acting on the qubits, the experimental setup must work in a way that H_1 can be applied for a precise interval of time, then turned off and followed by the application of H_2 of another interval of time and so on. Unfortunately, quantum gates cannot be applied perfectly; random errors are expected. Quantum error correction can remove the error generated from unreliable gates if the error is small enough.

5. A measurement capability specific for the qubit.

   The measurement on a qubit should give the outcome 0 with probability α and 1 with probability $1 - \alpha$, as explained in section 1.1, without changing the state of other qubits. This measurement is only ideal and is said to have 100% quantum efficiency. Usually this value is much lower than that. Some techniques have been developed to overcome this problem.

These requirements are not easy to be satisfied. The manipulation of qubits requires that they have to be accessible. However a quantum computer has to be well isolated to retain its quantum properties. For instance, a single nuclear spin has the advantage that superpositions on an external magnetic field can last for a long time; unfortunately, measuring the orientation of single nuclei can be challenging because their coupling to the world is minimal.

Some examples of physical realization of a quantum computer are the optical photon quantum computer [33], nuclear magnetic resonance (NMR) quantum computer [34] and superconducting qubits [35]; the latter is a promising and studied platform. In 2019 Google published a paper claiming to have achieved quantum supremacy [26].

Quantum algorithms are meant to be implemented in quantum computers where the physical error rate is below a certain threshold where, as stated by the Threshold theorem [36], errors can be suppressed by Quantum Error Correction (QEC) techniques. Unfortunately, these machines are still unavailable, and building an error-corrected quantum computer may take decades. The set of methods that allow qubits to be protected from quantum errors, introduced by decoherence and quantum noise, is called fault tolerant quantum computing. QEC methods, such as the Shor code that aims to correct bit flips and sign flips, requires ancilla qubits, which is not always possible in today's quantum computers. A method to reduce the error without ancilla qubits is called Quantum Error Mitigation (QEM); it can reduce the error in finding expectation values using classical post-processing and different runs of quantum circuits [11].

Originally proposed quantum algorithms incorporated with some QEC technique require millions of physical qubits, currently quantum computers are composed of about 100 qubits; these devices are often called Noisy Intermediate-Scale Quantum (NISQ) devices; they don't use QEC techniques. Therefore their operations are subject to errors, to overcome this problem, hybrid algorithms have been developed, where they partially run on a classical computer to deal with the decoherence errors of the NISQ devices. This approach is called near-term quantum computation. The NISQ era aims to exploit all the quantum computational power possible from these devices [11].

Chapter 2

Variational Quantum Algorithms

Quantum algorithms are algorithms that can be performed on a quantum computer. They are designed to outperform the efficiency of classical algorithms for some problem, redefining what is a tractable or intractable.
There are various models of quantum computation, but the most commonly used is the circuit model, that describes the computation in terms of a network of quantum gates, as seen in the previous chapter.

The first breakthrough in the development of quantum algorithms was the Shor's algorithm [22], designed to find the prime factors of an integer. Public-key cryptography, widely used in society today, relies on the fact that for classical computer factoring large integers is computationally intractable. If a quantum computer with enough qubits could operate without succumbing to quantum noise, then Shor's algorithm could be used to break public-key cryptography schemes.
Another example of a quantum algorithm that outperforms its classical counterpart is the Grover search algorithm [13] (or quantum search algorithm), the problem of locating an element in an unstructured list of N elements with a classical approach requires to check at least $N/2$ (N in the worst case) elements, instead with the Grover algorithm requires only about $\sqrt{2}$ steps.

2.1 Introduction to VQA

As introduced at the end of chapter 1.3, quantum computers available today are equipped with about 100 qubits where QEC methods are not applied. Often they are called Near Intermediate Scale Quantum (NISQ) devices and can implement the model of quantum circuits, but because of the noise each gate involves, the depth of circuits is very limited. Variational quantum algorithms (VQA) are a group of algorithms developed for NISQ devices. They delegate the classically difficult part of some computation to a quantum

computer and perform the classically tractable part on some sufficiently powerful classical device that optimizes the parameters of a parameterized quantum circuit $U(\vec{\theta})$. One of the two proposed VQA algorithms were the Variational Quantum Eigensolver (VQE) and the Quantum Approximate Optimization Algorithm (QAOA) [2], proposed to solve combinatorial optimization problems.

The VQA is composed by some modular components, shown in figure 2.1, mainly we have four components:

Objective function: the Hamiltonian $\mathcal{H}$ encodes information about a physical system, but also to problems not related to real physical systems, such as combinatorial problems. Writing a function in an operational form allow us to solve the problem encoded by the function using a quantum computer. The classical part of a VQA aims to minimize the cost function $\langle\Psi(\vec{\theta})|\,\mathcal{H}\,|\Psi(\vec{\theta})\rangle$, such that the ground state energy is calculated. [11].

Parameterized quantum circuits (PQC): Another essential constituent of a VQA is the quantum circuit that prepares the state that best meets the objective. It is generated through a unitary operation $U(\vec{\theta})$ that depends on some parameters $\vec{\theta}_{opt}$. Finding the parameter values $\vec{\theta}$ that deliver to an optimal $U(\vec{\theta}_{opt})$, on the other hand, is the task of the classical optimization subroutine. The unitary operation is applied to an initial state $|\psi\rangle_0$ that can be a superposition of the basis or a particular state motivated by the problem under consideration.

$$|\psi(\vec{\theta})\rangle = U(\vec{\theta})\,|\psi_0\rangle \tag{2.1}$$

Similarly, the initial parameters are unknown a priori. Usually they are chosen randomly, however knowledge about the problem can be exploited to choose a suitable initial parameters [11].

Measurement: In VQA, the action of gaining information about the quantum state is performed estimating the expectation value of the objective function $\langle\Psi(\vec{\theta})|\,\mathcal{H}\,|\Psi(\vec{\theta})\rangle$. The expectation value is estimated applying a unitary transformation in the diagonal basis of the observable $\mathcal{H}$. Is possible to obtain the expectation value directly from the probability of measuring specific computational states corresponding to an eigenvalue of $\mathcal{H}$. In a NISQ device, determining if a measured qubit is in the $|0\rangle$ or $|1\rangle$ state can be a expensive operation, because of the transformation to the diagonal basis. Alternatively the Pauli string representation can be used to efficiently parameterize the observable, as stated in 1.2.

Parameters optimization: The parameters of the PQC $U(\vec{\theta})$ are optimized by a

classical computer. This task is resolved by an optimization algorithm. It is important to consider that noisy data coming from current devices can invalidate the result of the optimization algorithm. This flaw makes that for PQC optimization there are better algorithms than others, and new algorithms are being proposed specifically for this purpose [11].

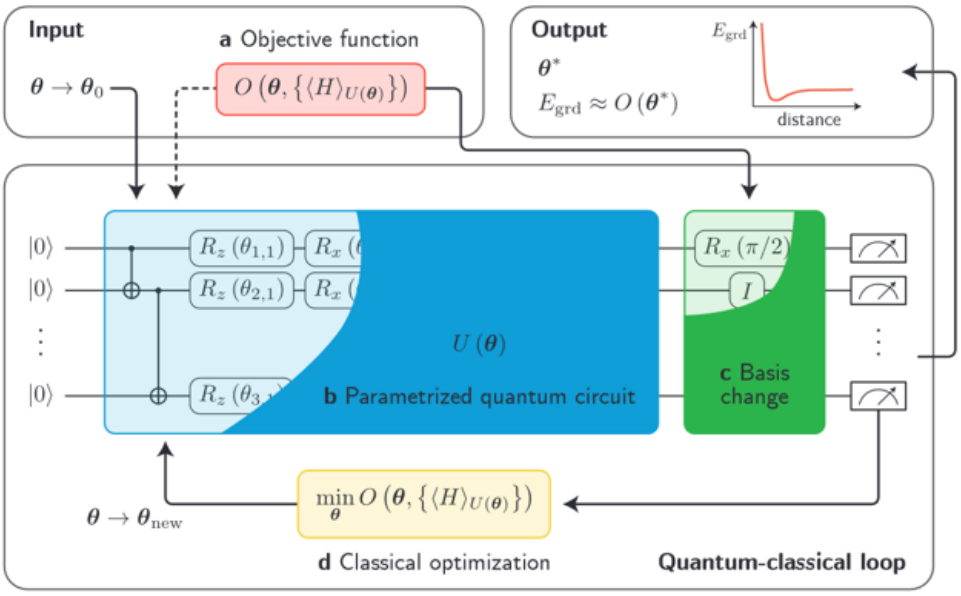

Figure 2.1: Representation of the VQA main components, *a)* the objective function O, *b)* the parameterized quantum circuit, *c)* the measurement scheme U, *d)* the classical optimizer that minimizes the objective. From [11]

2.2 The Quantum Approximate Optimization Algorithm

The QAOA (Quantum Approximate Optimization Algorithm), introduced by Farhi, Goldstone and Gutmann [2], is a variational quantum algorithm developed to solve combinatorial problems such as the Max-Cut or the Travelling salesman problems. These two problems are often used as benchmark in optimization methods because of their heavy computational effort in classical computation.

The Max-Cut problem is about the partition of the nodes of an undirected graph G in 2 groups such that the links between these groups are maximized. A variant of it is the Min-Cut, that aims to reduce the links between the 2 groups and can be solved by a classical computer with less computational effort than the Max-Cut, using the Edmonds and Karp algorithm [21]. The Traveling Salesman Problem tries to solve the problem of starting from a city (a node), finding the shortest route (weighted links) that cross each other city (the other nodes) only one time, and return to the original city.

The QAOA algorithm consists of the following steps:

1. Define a cost Hamiltonian H_c, such that its ground state encodes the solution to the optimization problem.

2. Define a mixer Hamiltonian H_m, such that $[H_c, H_m]=0$, usually chosen as $H_m = \sum_i^N X_i$, where N is the number of qubits and X_i is the X Pauli matrix applied to the qubit i.

3. Prepare an initial state $|\psi_0\rangle$, usually chosen as an equal superposition of the basis:

$$|\psi_0\rangle = \frac{1}{\sqrt{2^N}}(|0\rangle + |1\rangle)^{\otimes N} \tag{2.2}$$

4. Choose an integer $p \geq 1$, keeping in mind that the depth of the circuit grows (at worst) linearly with p, the quality of the approximation is expected to improve as p increases.

5. Initialize the parameters of the quantum circuit $\vec{\beta}_0$ and $\vec{\gamma}_0$ with dimension p, usually they are generated randomly in $[-2\pi, 2\pi]$. The parameter quantum circuit (PQC) is built by the repeated application of the two unitaries $e^{-i\beta_i H_C}$ and $e^{-i\gamma_i H_M}$, with $i = 1, ..., p$, in this way:

$$U(\vec{\beta}, \vec{\gamma}) = e^{-i\beta_p H_C} e^{-i\gamma_p H_M} \dots e^{-i\beta_2 H_C} e^{-i\gamma_2 H_M} e^{-i\beta_1 H_C} e^{-i\gamma_1 H_M} \tag{2.3}$$

6. Repeat the following loop until the convergence is reached:

   (a) The PQC is applied to the initial state:

$$|\psi\left(\vec{\beta}_{\mathrm{k}}, \vec{\gamma}_{\mathrm{k}}\right)\rangle = \mathrm{U}\left(\vec{\beta}_{\mathrm{k}}, \vec{\gamma}_{\mathrm{k}}\right)|\psi_0\rangle \tag{2.4}$$

   (b) The approximated solution of ground state of the Hamiltonian H_c is measured:

$$F_p = \langle\psi\left(\vec{\beta}_{\mathrm{k}}, \vec{\gamma}_{\mathrm{k}}\right)|\mathrm{H_C}|\psi\left(\vec{\beta}_{\mathrm{k}}, \vec{\gamma}_{\mathrm{k}}\right)\rangle \tag{2.5}$$

   (c) The value F_p is used by the classical optimizer to optimize the parameters of the parametrized quantum circuit $(\vec{\beta}_k, \vec{\gamma}_k) \Rightarrow (\vec{\beta}_{k+1}, \vec{\gamma}_{k+1})$ where $\mathrm{k} = 1, \ldots, \mathrm{N}$, and N is the total number of steps performed.

After the N steps are performed, $|\psi_{\mathrm{opt}}\rangle = |\psi\left(\vec{\beta}_{\mathrm{N}}, \vec{\gamma}_{\mathrm{N}}\right)\rangle$ is our approximation to the ground state of H_c, using this state is possible to calculate the approximation to the ground state $F_p = \langle\psi_{\mathrm{opt}}|H_C|\psi_{\mathrm{opt}}\rangle$, this value encodes the solution to our problem.

It has been estimated that for $p = 1$ and 420 qubits, it is possible to demonstrate the quantum advantage [37].

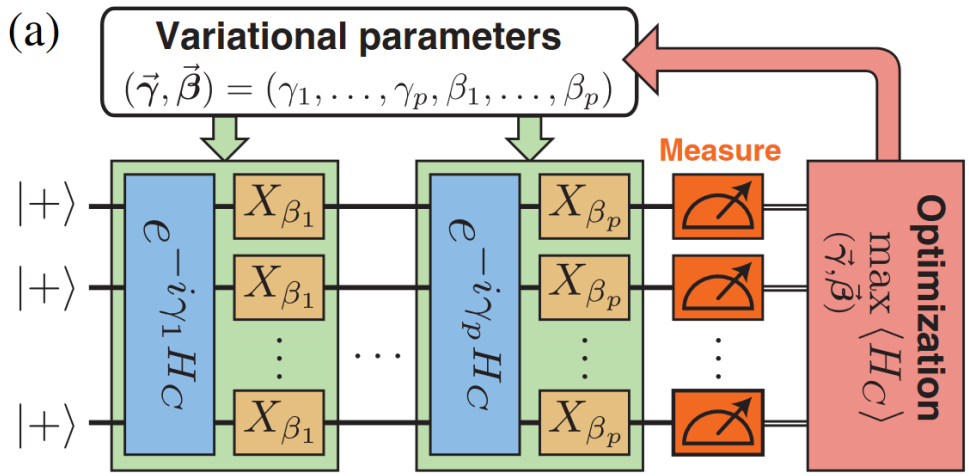

Figure 2.2: Scheme of the QAOA algorithm, the classical optimizer optimizes the parameters that characterize the quantum gates, the states are evolved and used to calculate an approximation of the ground state of the cost Hamiltonian; This algorithm is performed until the desired approximation is reached. From [12]

2.3 Optimization algorithms

Optimization algorithms are fundamental for variational quantum algorithms, but they are also crucial in many fields of science and engineering. Many problems in these fields can be solved as the optimization of some scalar parameterized objective function requiring maximization or minimization with respect to its parameters. Often they are gradient-based. The gradient-based optimization is a relatively efficient optimization method if the function is differentiable, since the computation of its first derivative is of the same computational complexity as the simple evaluation of the function [5].
The most known optimization algorithms gradient-based method is the stochastic gradient descent (SGD) but, with the increasing importance of these algorithms, extensions and variants have been developed, such as Adam optimizer.

The stochastic gradient descent

Stochastic gradient descent (SGD) [23] is an iterative method for optimizing an objective function $f(\vec{\theta})$ with suitable smoothness properties. In the beginning of the algorithm is selected a random point and at each iteration the SGD tries descend the slope updating the point until the minimum is reached.
The algorithm follows this steps:

1. Pick a random initial value for the parameters $\vec{\theta}_0$.
2. Calculate the gradient function $\nabla f(\vec{\theta}_0)$.
3. Calculate the step sizes for each dimension $\vec{w}_k := \vec{w}_k - \eta \nabla f(\vec{\theta}_k)$, where η is the stepsize (or learning rate), a parameter that can be optimized in order to improve the convergence.
4. Calculate the new parameters as: $\vec{\theta}_{k+1} := \vec{\theta}_k - \vec{w}_k$
5. Calculate the gradient with the new parameters $\nabla f(\vec{\theta}_{k+1})$
6. Repeat steps 3 to 5 for a reasonable number of steps or until gradient is almost 0.

Usually different values of the stepsize η are tested in order to improve the convergence. Smaller stepsizes makes the convergence very low, while larger stepsizes make the algorithm take huge steps down the slope, consequently it might jump across the minimum point and missing it. Usually, it is a good practise to set the stepsize to 0.01.

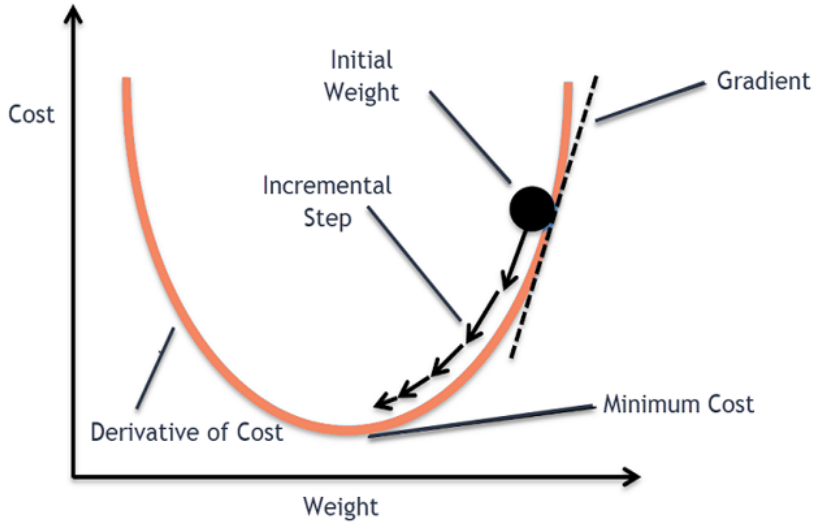

Figure 2.3: In the stochastic gradient descent the gradient of the cost function is calculated, for various steps, in order to minimize the latter. From [16]

The Adam optimization algorithm

The optimization algorithm Adam (P. Kingma & Lei Ba, 2015 [5]), which derived his name from adaptive moment estimation, is a method for efficient stochastic optimization that only requires first-order gradients. The method can adapt individually the learning rates for different parameters from estimates of first and second moments of the gradients. The Adam was developed combining the advantages of two others optimizer: RMSProp (Tieleman & Hinton, 2012 [25]), an algorithm designed for artificial neural network training, and AdaGrad (Duchi et al., 2011 [24]), which scale the learning rate for each dimension. An important advantage of the Adam method is that the magnitudes of parameter updates are invariant to the gradient rescaling. The Adam optimizer depends on these parameters: α is the stepsize, similarly at the stochastic gradient descent, it controls the movement in the search space , β_1 and β_2 are the exponential decay rates for the 1° and the 2° moment estimates, ϵ is a small constant for numerical stability.
Let $f(\vec{\theta})$ be the objective function, the following is the pseudo-code of the Adam taken from [5]:

Require: α: Stepsize
Require: $\beta_1,\ \beta_2 \in [0, 1)$: Exponential decay rates for the moment estimates
Require: $f(\theta)$: Stochastic objective function with parameters θ
Require: θ_0: Initial parameter vector
$m_0 \leftarrow 0$ (Initialize 1st moment vector)
$v_0 \leftarrow 0$ (Initialize 2nd moment vector)
$t \leftarrow 0$ (Initialize timestep)
while θ_t not converged **do**
$t \leftarrow t + 1$
$g_t \leftarrow \nabla_\theta f_t(\theta_{t-1})$ (Get gradients w.r.t. stochastic objective at timestep t)
$m_t \leftarrow \beta_1 \cdot m_{t-1} + (1 - \beta_1) \cdot g_t$ (Update biased first moment estimate)
$v_t \leftarrow \beta_2 \cdot v_{t-1} + (1 - \beta_2) \cdot g_t^2$ (Update biased second raw moment estimate)
$\hat{m}_t \leftarrow m_t/(1 - \beta_1^t)$ (Compute bias-corrected first moment estimate)
$\hat{v}_t \leftarrow v_t/(1 - \beta_2^t)$ (Compute bias-corrected second raw moment estimate)
$\theta_t \leftarrow \theta_{t-1} - \alpha \cdot \hat{m}_t/(\sqrt{\hat{v}_t} + \epsilon)$ (Update parameters)
end while
return θ_t (Resulting parameters)

Good default settings are $\alpha = 0.001$, $\beta_1 = 0.9$, $\beta_2 = 0.999$ and $\epsilon = 10^{-8}$.

The BFGS algorithm

The Broyden, Fletcher, Goldfarb, and Shanno (or BFGS) algorithm [14], is one of the most used second-order algorithms for numerical optimization. It is principally used for the logistic regression algorithm in machine learning.

It makes use of the second-order derivative of an objective function, therefore is referred as second-order optimization algorithm. It belongs to a class of algorithms called Quasi-Newton methods that approximates the second derivative (the Hessian) for optimization problems, important, especially where the second derivative cannot be calculated.

The Newton's method work in this way, a function $f(x_k)$ is approximated at the point x_k with a paraboloid, and then the approximation is minimized stepping to the minimum of that paraboloid. Differently to the stochastic gradient descent introduced before, there is no longer a need to set a learning rate parameter η, in the BFGS the step size is determined by the distance from x_k to the minimum of the fitted parabola. The Newtonian method has mainly two problems, is too sensitive to initial conditions and can be very computationally expensive in proportion to the number of parameters. The Quasi-Newtonian method was developed to overcome these problems, it takes the best of the SGD method and Newton's method, it has a faster convergence than the SGD, and a

lower cost per iteration than Newton's method [15].

In the BFGS method, the descent direction is determined by preconditioning the gradient with curvature information. It does so by gradually improving an approximation to the Hessian matrix of the cost function.
In Newton's method, the Hessian of the function are computed, instead in quasi-Newtonian method, we just approximate it with a matrix B_k, which is updated iteration after iteration using information computed from previous steps.
The search direction p_k at step k is given a the solution that is similar to the Newton equation:

$$B_k \mathbf{p}_k = -\nabla f(\mathbf{x}_k) \tag{2.6}$$

where B_k is an approximation to the Hessian matrix, and $f(\mathbf{x}_k)$ is the function to minimize. The search direction characterized by $\mathbf{p}_k$ is then used to find the next step $\mathbf{x}_{k+1}$ by minimizing $f(\mathbf{x}_k + \gamma \mathbf{p}_k)$ where $\gamma > 0$.
The quasi-Newton condition imposed on the update of B_k is

$$B_{k+1}\left(\mathbf{x}_{k+1} - \mathbf{x}_k\right) = \nabla f\left(\mathbf{x}_{k+1}\right) - \nabla f\left(\mathbf{x}_k\right) \tag{2.7}$$

At the point x_{k+1} is computed the approximation to the Hessian matrix B_{k+1}, and at the approximate Hessian, at step k, is summed to two matrices:

$$B_{k+1} = B_k + U_k + V_k \tag{2.8}$$

Where U_k and V_k are symmetric rank-one matrices, and their sum is a rank-two update matrix. To maintain the symmetry and positive definiteness of B_{k+1}, the update form can be chosen as

$$B_{k+1} = B_k + \alpha \mathbf{u}\mathbf{u}^\top + \beta \mathbf{v}\mathbf{v}^\top \tag{2.9}$$

Imposing the secant condition, $B_{k+1}\mathbf{s}_k = \mathbf{y}_k$. Choosing $\mathbf{u} = \mathbf{y}_k$ and $\mathbf{v} = B_k\mathbf{s}_k$, we can obtain:

$$\alpha = \frac{1}{\mathbf{y}_k^T \mathbf{s}_k} \quad \text{and} \quad \beta = -\frac{1}{\mathbf{s}_k^T B_k \mathbf{s}_k} \tag{2.10}$$

Finally, we substitute α and β into $B_{k+1} = B_k + \alpha \mathbf{u}\mathbf{u}^\top + \beta \mathbf{v}\mathbf{v}^\top$ and get the update equation of B_{k+1}:

$$B_{k+1} = B_k + \frac{\mathbf{y}_k \mathbf{y}_k^{\mathrm{T}}}{\mathbf{y}_k^{\mathrm{T}} \mathbf{s}_k} - \frac{B_k \mathbf{s}_k \mathbf{s}_k^{\mathrm{T}} B_k^{\mathrm{T}}}{\mathbf{s}_k^{\mathrm{T}} B_k \mathbf{s}_k} \tag{2.11}$$

The algorithm is the following:

1. Obtain a direction $\mathbf{p}_k = -\nabla f(\mathbf{x}_k) B_k^{-1}$
2. Perform a one-dimensional optimization to estimate an acceptable stepsize α_k in the direction found in the first step.
3. Set $\mathbf{s}_k = \alpha_k \mathbf{p}_k$ and update $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{s}_k$
4. Calculate the difference $\mathbf{y}_k = \nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k)$
5. Calculate $B_{k+1} = B_k + \frac{\mathbf{y}_k \mathbf{y}_k^{\mathrm{T}}}{\mathbf{y}_k^{\mathrm{T}} \mathbf{s}_k} - \frac{B_k \mathbf{s}_k \mathbf{s}_k^{\mathrm{T}} B_k^{\mathrm{T}}}{\mathbf{s}_k^{\mathrm{T}} B_k \mathbf{s}_k}$ and go back to step 1

A possible inizialization for B_k is $B_0 = \mathbb{1}$, in this case the first step will be equivalent to a gradient descent.
The first step of the algorithm is performed using the inverse of the matrix B_k.

$$B_{k+1}^{-1} = \left(I - \frac{\mathbf{s}_k \mathbf{y}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \right) B_k^{-1} \left(I - \frac{\mathbf{y}_k \mathbf{s}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \right) + \frac{\mathbf{s}_k \mathbf{s}_k^T}{\mathbf{y}_k^T \mathbf{s}_k} \tag{2.12}$$

The step 5 can be rewritten such that B_k must be the approximation of the inverse of the Hessian instead of the Hessian approximation to avoid matrix inversion.

Basin-hopping optimization

The basin-hopping algorithm, by Wales et al. [7], is a stochastic global optimization method that attempts to find the global minimum of an objective function. It has been motivated by problems from physical chemistry, especially to find configurations with the lowest energy or stable molecular configurations. These tasks are characterized by a lots of local minima which makes it hard for standard optimization methods because of the very strong dependency on the initial conditions.

The basin-hopping is composed by two ingredients, the Monte Carlo approach, and a local optimizer that returns for a given starting point a local minimum.

The algorithm is composed by the following steps:

1. Choose a starting point.
2. Compute a local minimum with the associated local optimizer.
3. Apply a random perturbation on the coordinates of the local minimum.
4. Compute the next local minimum.
5. Compare the local minima using a criterion previously chosen and keep the best.

The random displacement obviously needs to be sufficiently large to escape the current local minimum but no too large that the hopping doesn't become completely random.

The step for the random displacement is bounded by the *stepsize*, an important basin-hopping parameter, that bounds the size of the step. The step is chosen uniformly in the region from $\vec{x_0}$-*stepsize* to $\vec{x_0}$+*stepsize* where $\vec{x_0}$ is the position of a local minimum previously reached. Ideally, it should be comparable to the typical separation (in argument values) between local minima of the function being optimized. This method, by default, adjusts the stepsize to find an optimal value, but this may take many iterations, the algorithm can be faster if a sensible initial value for stepsize is initially chosen.

A possible criterion for the comparison between the minima is the Metropolis criterion, where the new local minimum is always chosen if $f(\mathbf{x}_{new}) < f(\mathbf{x}_{old})$, if this is not the case, the new local minimum is chosen with probability

$$e^{-(f(\mathbf{x}_{new})-f(\mathbf{x}_{old}))/T} \tag{2.13}$$

Where T, the *temperature*, is a parameter to optimize. To get better results, T should be comparable to the typical difference (in function values) between local minima. If $T = 0$, the algorithm is called "Monotonic basin-hopping", in which all steps that increase energy are not taken in consideration [8].

A lot of molecular systems have been optimized primarily using basin-hopping, their are stored in the Cambridge Cluster Database [6], this database includes minimization problems exceeding 300 degrees of freedom.

Chapter 3

Diatomic molecules

We need a Hamiltonian to benchmark the algorithm for the calculation of the excited states using the QAOA. The diatomic molecule Hamiltonian will be used for this purpose.

This chapter will introduce the properties of diatomic molecules, starting from the study of the Hydrogen atom and the H_2^+ molecule. Only for one-electron systems, the exact solution of the Schrodinger equation can be calculated, for other systems approximations are necessary, such as the LCAO approximation. Usually, in computational chemistry, one-electron functions are approximated using a finite basis set. There are many basis sets, where the choice depends on the physics of the problem.

In the end, we will see the limits of classical computation applied to the problems of chemistry, and how quantum computation can help to overcome these limits. An electronic system can be treated by a quantum computer mapping the orbitals by qubits. This is achieved using the electronic Hamiltonian in second quantization, where the creation and annihilation operators are mapped into qubits using a transformation such as the Jordan-Wigner transformation.
The main reference used for this part is [3]

3.1 The Hydrogen atom

The hydrogen atom is the simplest atom in nature and a good starting point for studying atomic structure. It consists of an electron moving in the proton's spherically symmetric Coulomb potential. Hydrogen and other one-electron systems are the only real systems for which the Schrödinger equation can be solved exactly. For all other atoms or molecules, approximations must be made.

3.1.1 Schrödinger Equation for one-electron systems

The Schrödinger equation can be viewed as the fundamental equation of non-relativistic quantum theory. The treatment of the hydrogen atom can explain the basic ideas of quantum mechanics.

The Schrödinger equation for a system consisting of one electron (mass m_1, charge $q = -e$ and radius vector $\boldsymbol{r}_1$) and a nucleus (mass $m_2 \gg m_1$, charge $q = +Ze$ and radius vector $\boldsymbol{r}_2$) is:

$$-\frac{\hbar^2}{2m_1}\Delta_1\psi - \frac{\hbar^2}{2m_2}\Delta_2\psi - \frac{Ze^2}{4\pi\varepsilon_0 r}\psi = E\psi\left(\boldsymbol{r}_1, \boldsymbol{r}_2\right), \tag{3.1}$$

where Δ_i is the Laplace operator respect to r_i. The first term describes the kinetic energy of the electron, the second the kinetic energy for the nucleus and the third the Coulomb interaction between the two particles, where $r = |r_1 - r_2|$.

A closed particle system can always be broken down into the movement of the center of mass and the relative movement of the particles in the center of mass:

$$\boldsymbol{R} = \frac{m_1\boldsymbol{r}_1 + m_2\boldsymbol{r}_2}{M} \quad \text{with} \quad M = m_1 + m_2, \quad \boldsymbol{R} = \{X, Y, Z\}.$$

and with relative distance $\boldsymbol{r} = \{x, y, z\} = |r_1 - r_2|$. Using standard techniques, we use separation of variables by setting $\psi(\boldsymbol{R}, \boldsymbol{r}) = f(\boldsymbol{r})g(\boldsymbol{R})$ to obtain an equation for the relative motion:

$$-\frac{\hbar^2}{2\mu}\Delta\psi + E_{\text{pot}}\left(r\right)\psi = E\psi \tag{3.2}$$

that is the Schrodinger equation for a particle in a spherically symmetric potential: $E_{pot}(r) = -2e^2/4\pi\epsilon_0 r$.

3.1.2 Quantum Numbers

For any spherical potential, the wave function (3.2) can be broken down into a radial function $R_{n,la}(r)$, dependent on distance proton-electron r, on the principal quantum number n and orbital angular quantum number l, and spherical functions Y_l^m, dependent from the angular components (θ, ϕ), on the magnetic quantum number m and the orbital angular quantum number l, the wave function can be written as:

$$\psi_{n,l,m}(r, \vartheta, \varphi) \propto R_{n,l}(r)Y_l^m(\vartheta, \varphi) \tag{3.3}$$

are also called atomic orbitals. The normalized total wave functions ψ for the lowest energy states of the hydrogen atom are compiled in Table 3.2.

l	state label	$\lvert m\rvert$	state label
0	s	0	σ
1	p	1	π
2	d	2	δ
3	f	3	φ
4	g	4	γ

Table 3.1: Labeling of atomic states (l, m) using Latin and Greek letters

n	l	m	Eigenfunction $\psi_{n,l,m}(r, \vartheta, \varphi)$
1	0	0	$\frac{1}{\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \mathrm{e}^{-Zr/a_0}$
2	0	0	$\frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \left(2 - \frac{Zr}{a_0}\right) \mathrm{e}^{-Zr/2a_0}$
2	1	0	$\frac{1}{4\sqrt{2\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \frac{\mathrm{Zr}}{a_0} \mathrm{e}^{-Zr/2a_0} \cos\vartheta$
2	1	± 1	$\frac{1}{8\sqrt{\pi}} \left(\frac{Z}{a_0}\right)^{3/2} \frac{Zr}{a_0} \mathrm{e}^{-Zr/2a_0} \sin\vartheta \mathrm{e}^{\pm i\varphi}$

Table 3.2: Normalized total wave functions of an electron in the Coulomb potential $E_{\mathrm{pot}} = -Ze^2/(4\pi\varepsilon_0 r)$

Each atomic state, described by its energy and spatial distribution of electrons, is uniquely defined by the three quantum numbers n, l and m, summarized in Table 3.1.

The energy E_n of an atomic state in this model depends exclusively on n, therefore, all states with possible combinations of l and m, for a fixed value of n, have the same energy. Not all sets of quantum numbers (n, l, m) are possible. For example, the orbital angular quantum number l can never be greater or equal to the principal quantum number $n (l < n)$. They must obey the constraints:

- $n = 1, 2, 3, \ldots$
- $l = 0, 1, 2, \ldots, (n-1)$
- $m = -l, (-l+1), \ldots, 0, \ldots, (+l-1), +l$

As a consequence for the magnetic quantum number m there are $2l + 1$ possible values, because $-l \leq m \leq +l$.

The radial wave functions of some atomic states are shown in Fig. 3.1. Different states with the same energy are called degenerate states. The number of degenerate states is

called the degeneracy order.

We must be careful that several effects, such as electron spin, nuclear spins and external fields, not included in the Schrödinger theory. They can break the degeneracy and split degenerate levels into components with different energies.

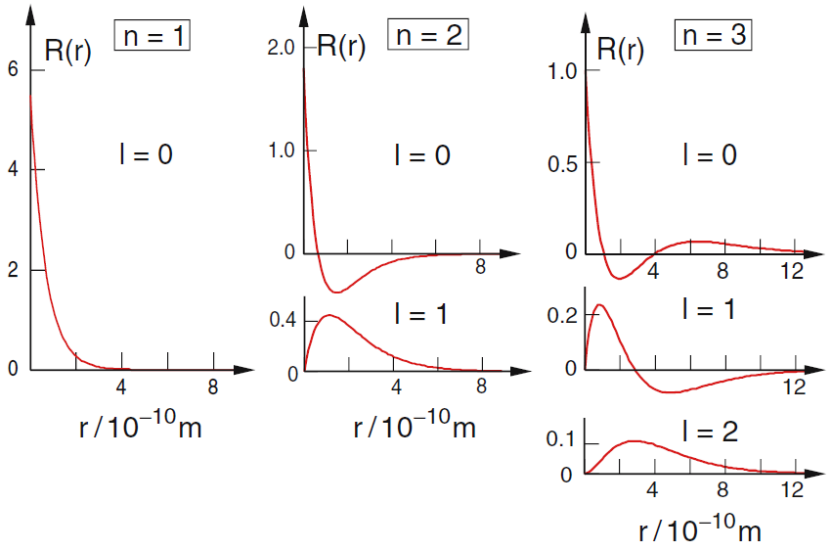

Figure 3.1: The radial wave function $R_{n,l}(r)$ for the principle quantum numbers $n = 1, 2, 3$. From [3]

3.2 The H_2^+ molecular ion

The H_2^+ ion, which has two identical nuclei (protons) and one electron, is the most basic of all molecules. The interaction potential $\mathrm{E_{pot}}$ between the three particles is

$$E_{\mathrm{pot}} = -\frac{e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_{\mathrm{A}}} + \frac{1}{r_{\mathrm{B}}} - \frac{1}{R}\right) \tag{3.4}$$

where R is the distance between the 2 protons, and r_A and r_B are distances from the proton A and B to the electron, as showed in Fig. 3.2

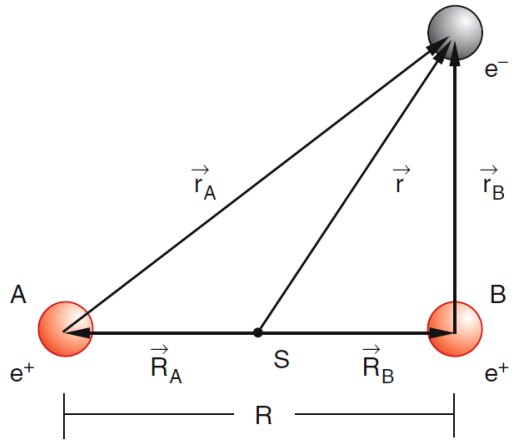

Figure 3.2: The H_2^+ molecule. From [3]

When we choose the origin of our coordinate system as the center of mass of the two protons (the small mass the electron does not noticeably change the center of mass) we obtain from Fig. 3.2 the relations

$$\boldsymbol{r} = \boldsymbol{R}_{\mathrm{A}} + \boldsymbol{r}_{\mathrm{A}} = \boldsymbol{R}_{\mathrm{B}} + \boldsymbol{r}_{\mathrm{B}} \Rightarrow \boldsymbol{r} = \frac{1}{2}\left(\boldsymbol{r}_{\mathrm{A}} + \boldsymbol{r}_{\mathrm{B}}\right) \tag{3.5}$$

because $\boldsymbol{R}_{\mathrm{A}} = -\boldsymbol{R}_{\mathrm{B}}$. With $\boldsymbol{R} = \boldsymbol{r}_{\mathrm{A}} - \boldsymbol{r}_{\mathrm{B}}$ we can replace $\boldsymbol{r}_{\mathrm{A}}$ and $\boldsymbol{r}_{\mathrm{B}}$ by:

$$\boldsymbol{r}_{\mathrm{A}} = \boldsymbol{r} + \frac{1}{2}\boldsymbol{R}; \quad \boldsymbol{r}_{\mathrm{B}} = \boldsymbol{r} - \frac{1}{2}\boldsymbol{R} \tag{3.6}$$

The Schrödinger equation for this three-body problem is:

$$\hat{H}\Phi = E\Phi \tag{3.7}$$

with

$$\hat{H} = -\frac{\hbar^2}{2M}\left(\Delta_{\mathrm{A}}\left(\boldsymbol{R}_{\mathrm{A}}\right) + \Delta_{\mathrm{B}}\left(\boldsymbol{R}_{\mathrm{B}}\right)\right) - \frac{\hbar^2}{2m}\Delta_e(\boldsymbol{r}) + E_{\mathrm{pot}}(\boldsymbol{r}, \boldsymbol{R}) \tag{3.8}$$

where the kinetic energy of the nuclei is represented by the first two terms, and that of the electron by the third. The nuclei and the electron move in the potential $E_{\text{pot}}(\boldsymbol{r}, R)$.

3.2.1 A Solution for the H_2^+ Molecule

In contrast to the case for the H atom, the Schrödinger equation cannot be analytically solved. Approximations must be taken. Calling M the mass of the proton and m the mass of the electron, we can ignore the kinetic energy of the nuclei $E_{\text{kin}} = p^2/2M$, because $(M/m \approx 1836)$, considering the nuclei at fixed distance R (Born-Oppenheimer approximation) [3].
For each selected value of R, we have the wave function, the electron's energy, and the repulsive potential energy between the two nuclei. It follows that all pertinent values of R are resolved for the equation (3.7).
In this approximation of the rigid H_2^+ the Schrödinger equation (3.7), making explicit the potential (3.4), becomes:

$$\left[-\frac{\hbar^2}{2m}\Delta_e(\boldsymbol{r}) - \frac{e^2}{4\pi\varepsilon_0}\left(\frac{1}{r_{\text{A}}} + \frac{1}{r_{\text{B}}} - \frac{1}{R}\right)\right]\Phi\left(\boldsymbol{r}_{\text{A}}, \boldsymbol{r}_{\text{B}}, R\right) = E(R)\Phi\left(\boldsymbol{r}_{\text{A}}, \boldsymbol{r}_{\text{B}}, R\right) \tag{3.9}$$

where $\boldsymbol{r}_{\text{A}}$ and $\boldsymbol{r}_{\text{B}}$ depend on the coordinates both of the electron and of the two nuclei. Equation (3.9) can be solved analytically in a similar way as (3.1) for the H atom. It is convenient to use elliptical coordinates because the potential energy no longer has spherical symmetry but instead cylindrical symmetry. The line linking the two nuclei is the symmetry axis (z-axis), and the two nuclei are the focal points of an ellipsoid with cylindrical symmetry.
Following the proof in [3], we have that the wave functions result in discrete energy eigenvalues $E_n(R)$ and depend on the internuclear separation R (Fig. 3.3).
The functions $E(R)$ can either exhibit minima at a specific value of R if the energy states are stable or decrease inversely with R (repulsive instable energy states).
These curves

$$E(R) = \left\langle E_{\text{kin}}\left(\text{e}^-\right)\right\rangle + \frac{e^2}{4\pi\varepsilon_0}\left(\frac{1}{R} - \left\langle\frac{1}{r_{\text{A}}} + \frac{1}{r_{\text{B}}}\right\rangle\right) \tag{3.10}$$

are called potential energy curves, although they also include the averaged kinetic energy of the electron.
The angular momentum l of the electron is no longer constant because the potential is not spherically symmetric. It precesses around the internuclear axis. Its absolute value $|l|$ depends on the internuclear separation R, but its projection onto the z-axis l_z has a well-defined expectation value:

$$\langle l_z\rangle = m\hbar, \tag{3.11a}$$

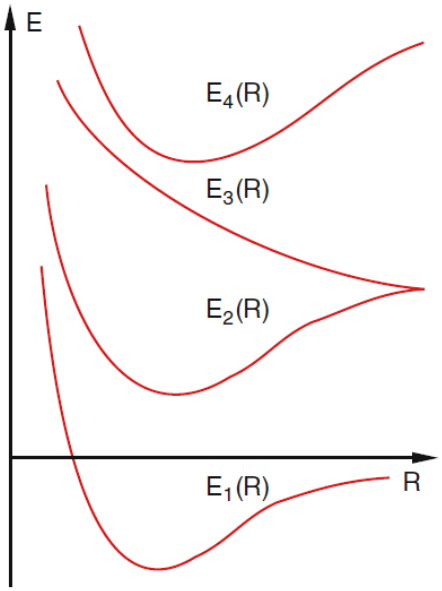

Figure 3.3: The potential curves $E(R)$ for the H_2^+ molecule, corresponding to bound states and repulsive unstable states. From [3]

which is determined by the integer $m = 0, \pm 1, \pm 2, \ldots, \pm l$ (Fig. 3.4) and is, for a given potential curve $E(R)$, independent of R, because is only dependent on φ.

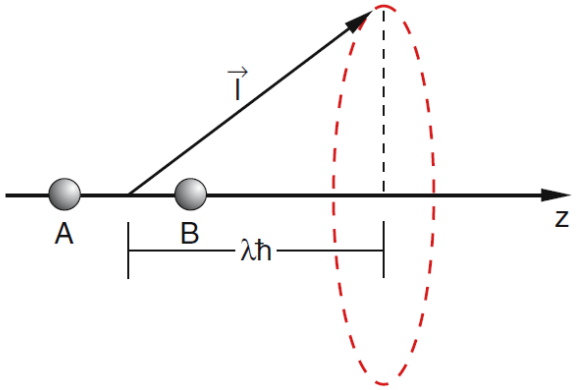

Figure 3.4: The orbital angular momentum l of the electron that precess on the internuclear axis and its constant projection $l_z = \lambda$. From [3]

The energy of a level does not depend on the direction of the field in the axial electric field of a non-rotating diatomic molecule. This indicates that levels with the same energy have the following coordinates: $\langle l_z \rangle = \pm m \cdot \hbar$, we have that the quantum number $\lambda = |m|$ describes the molecular levels. Instead of (3.11a) we write

$$\langle l_z \rangle = \lambda\hbar \tag{3.11b}$$

The electrons in diatomic molecules are called σ-electrons for $\lambda = 0$; π electrons for $\lambda = 1$ and δ-electrons for $\lambda = 2$ etc.

A magnetic field in the z direction is created as a result of the electron's precession around the internuclear axis for $\lambda > 0$. The magnetic moment μ_s can have two different orientations in this field if the electron spin s is taken into account, which is similar to the situation for atoms in the Stern-Gerlach experiment. Only the z-component of the electron spin precesses around the magnetic field direction, which is the z-direction. The expectation value for the z-component electron spin

$$\langle s_z \rangle = m_s \cdot \hbar = \pm\frac{1}{2}\hbar \tag{3.11c}$$

has definite eigenvalues (Fig. 3.5).

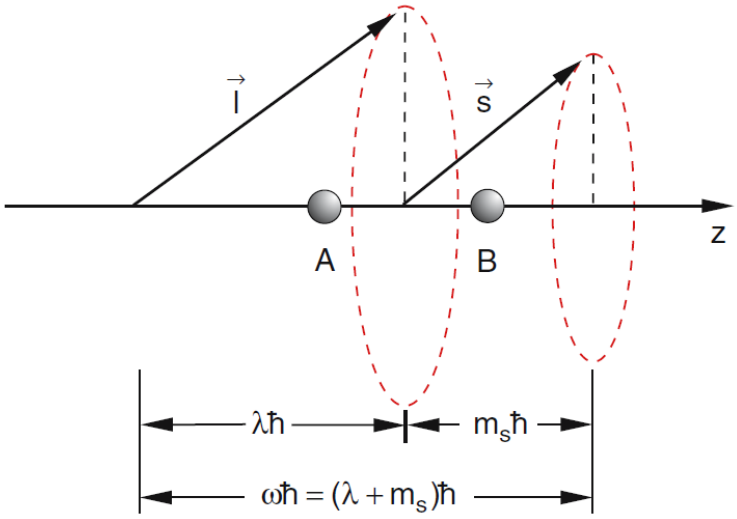

Figure 3.5: The projections of the orbital angular momentum l and the electron spin s adds up to the total projection $\omega\hbar = (\lambda + m_s)\hbar$. From [3]

3.2.2 The LCAO Approximation

The rigid H_2^+ molecule can be rigorously treated, but it is instructive to use some approximation methods to describe the molecule and compare the results with the exact solutions, and learn more about the advantages and disadvantages of these methods [3].

One of these techniques consists into approximating the molecular wavefunction as a linear combination of the atomic orbitals (LCAO) of the constituent atoms. The coefficients of this linear combination are chosen to minimize the energy calculated through the molecular wave function. The correct wave function creates always lower energies than the approximative ones.
The absolute square of the molecular orbital wave function $|\Phi(x,y,z)|^2$ represents the spatial probability density distribution of the electron and corresponds to the classical orbitals of the electron in the molecule.
The H_2^+ molecular ion can be composed of a H atom (nucleus A) and a H^+-ion (nucleus B), where the lowest energy state of H_2^+ corresponds to the H atom in its $1s$ ground state. The atomic orbital of the electron in the H atom is then (see table 3.2):

$$\phi_{\mathrm{A}}\left(r_{\mathrm{A}}\right) = \frac{1}{\sqrt{\pi a_0^3}} \mathrm{e}^{-r_{\mathrm{A}}/a_0}. \tag{3.12}$$

where r_a is the distance between the electron and the nucleus A. Either nucleus A or B can have the electron nearby, both possibilities must be considered because we are unable to distinguish them. As a result we assume:

$$\Phi(\boldsymbol{r}, R) = c_1 \phi_{\mathrm{A}}\left(\boldsymbol{r}_{\mathrm{A}}\right) + c_2 \phi_{\mathrm{B}}\left(\boldsymbol{r}_{\mathrm{B}}\right) \tag{3.13}$$

where $\boldsymbol{r}_{\mathrm{A}}$ and $\boldsymbol{r}_{\mathrm{B}}$ can be substituted by the internuclear distance R and the distance from the electron to the center of mass $r = |\boldsymbol{r}|$ (Fig. 3.2).
The wave function should be normalized for arbitrary values of R. This demands the integration is performed over the coordinates of the electron.

$$\int |\Phi|^2 \, \mathrm{d}\boldsymbol{r} = c_1^2 \int \left|\phi_{\mathrm{A}}\left(\boldsymbol{r}_{\mathrm{A}}\right)\right|^2 \, \mathrm{d}\boldsymbol{r}_{\mathrm{A}} + c_2^2 \int \left|\phi_{\mathrm{B}}\left(\boldsymbol{r}_{\mathrm{B}}\right)\right|^2 \, \mathrm{d}\boldsymbol{r}_{\mathrm{B}} + 2c_1 c_2 \int \phi_{\mathrm{A}} \phi_{\mathrm{B}} \mathrm{d}^3 r = 1 \tag{3.14}$$

The first two integrals go to 1 because the atomic orbitals are already normalized, instead for the coefficients, from (3.14) we obtain the condition for the normalized molecular wave function:

$$c_1^2 + c_2^2 + 2c_1 c_2 S_{\mathrm{AB}} = 1 \tag{3.15}$$

where S_{AB} is called overlap integral and is given by the integral:

$$S_{\mathrm{AB}} = \int \phi_{\mathrm{A}}\left(\boldsymbol{r}_{\mathrm{A}}\right) \phi_{\mathrm{B}}\left(\boldsymbol{r}_{\mathrm{B}}\right) \mathrm{d}\tau \quad \text{with} \quad \mathrm{d}\tau = r^2 \sin\vartheta \mathrm{d}r \mathrm{d}\vartheta \mathrm{d}\varphi, \tag{3.16}$$

Because the integration is over the electron coordinates $\boldsymbol{r}$, which depend on R, its value is dependent on this parameter, as a consequence of symmetries, we have the condition $|c_1|^2 = |c_2|^2 = |c|^2$, additionally, the exchange of the two atomic orbitals

requires the molecular wave function to be either symmetric or antisymmetric. The normalized molecular wave functions are therefore given by:

$$\Phi_{\mathrm{s}} = \frac{1}{\sqrt{2+2S_{\mathrm{AB}}}}\left(\phi_{\mathrm{A}} + \phi_{\mathrm{B}}\right) \quad \text{and } \Phi_{\mathrm{a}} = \frac{1}{\sqrt{2-2S_{\mathrm{AB}}}}\left(\phi_{\mathrm{A}} - \phi_{\mathrm{B}}\right) \tag{3.17}$$

as shown by Fig. 3.6:

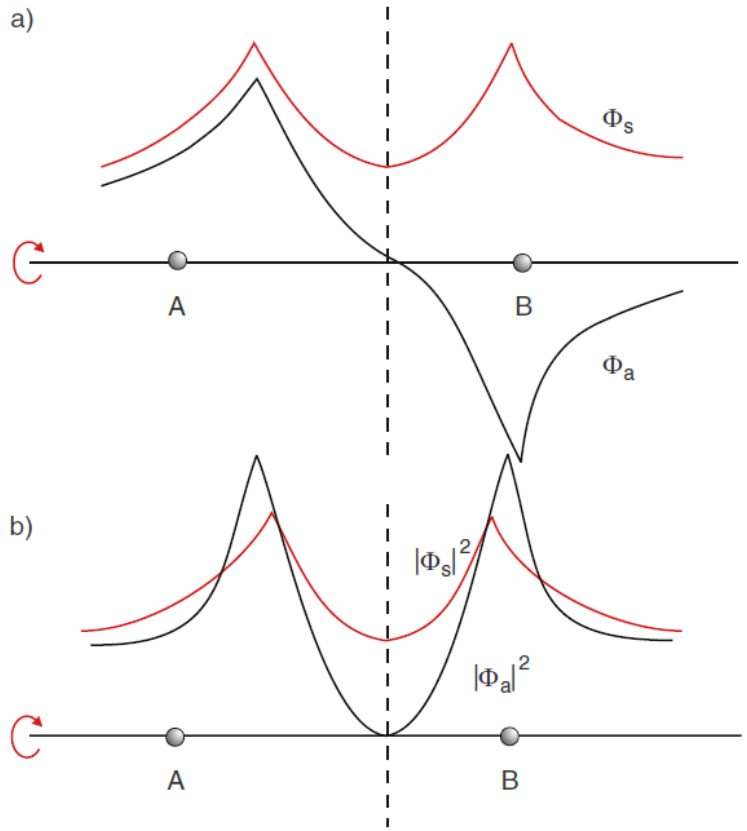

Figure 3.6: Cut through the surfaces with cylindrical symmetry (a) ψ_s and ψ_a (b) $|\Phi_s|^2$ and $|\Phi_a|^2$. From [3]

The expectation value for the energy is

$$\langle E \rangle = \int \Phi \hat{H} \Phi \mathrm{d}\tau \tag{3.18}$$

where $\hat{H}$ is the Hamiltonian in the Schrödinger equation (3.9) of the rigid molecule. Inserting (3.18) with (3.17) into the Schrödinger gives two energies

$$E_{\mathrm{s}}(R) = \frac{H_{\mathrm{AA}} + H_{\mathrm{AB}}}{1 + S_{\mathrm{AB}}} \quad \text{and}$$
$$E_{\mathrm{a}}(R) = \frac{H_{\mathrm{AA}} - H_{\mathrm{A0}}}{1 - S_{\mathrm{AB}}} \tag{3.19}$$

which depend on the nuclear distance R, because the following integrals

$$\begin{aligned} H_{\mathrm{AA}} &= \int \phi_{\mathrm{A}} \hat{H} \phi_{\mathrm{A}} \mathrm{d}\tau \\ H_{\mathrm{BB}} &= \int \phi_{\mathrm{B}} \hat{H} \phi_{\mathrm{B}} \mathrm{d}\tau \\ H_{\mathrm{AB}} &= \int \phi_{\mathrm{A}} \hat{H} \phi_{\mathrm{B}} \mathrm{d}\tau \quad \text{and} \\ S_{\mathrm{AB}} &= \int \phi_{\mathrm{A}} \phi_{\mathrm{B}} \mathrm{d}\tau \end{aligned} \tag{3.20}$$

depend on R.
The Fig. 3.7 shows the potential curves $E_{\mathrm{s}}(R)$ and $E_{\mathrm{a}}(R)$. While $E_{\mathrm{s}}(R)$ has a minimum and corresponds to a bound state, $E_{\mathrm{a}}(R)$ falls with increasing R. It corresponds to an unstable molecular state and represents a repulsive potential.

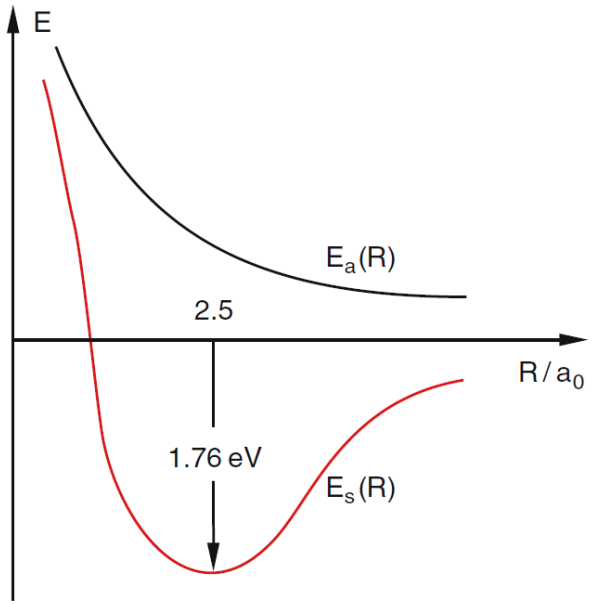

Figure 3.7: Potential curves $E_a(R)$ for the antisymmetric distribution $|\phi_a|^2$ and for $E_s(R)$ the symmetric charge distribution $|\phi_s|^2$ calculated in the LCAO approximation. From [3]

The LCAO approximation in its simplest form is not in good agreement with the correct solutions, this happens because the LCAO approximation doesn't describe accurately the following effects:

- The electron charge distribution, having its maximum in the middle between the two nuclei pulls the two nuclei towards each other.
- The molecular wave function Φ_s has a larger spatial extension than the atomic 1s orbitals. The spatial uncertainty for the electron is increased and its momentum uncertainty is therefore smaller than for the H-atom, leading to a decrease in the kinetic energy,

The approximation must be improved.

3.2.3 The H_2 molecule

The H_2 molecule (Fig. 3.8) consists of two protons and two electrons, whose interactions, differently from the H_2^+, must be taken in account.

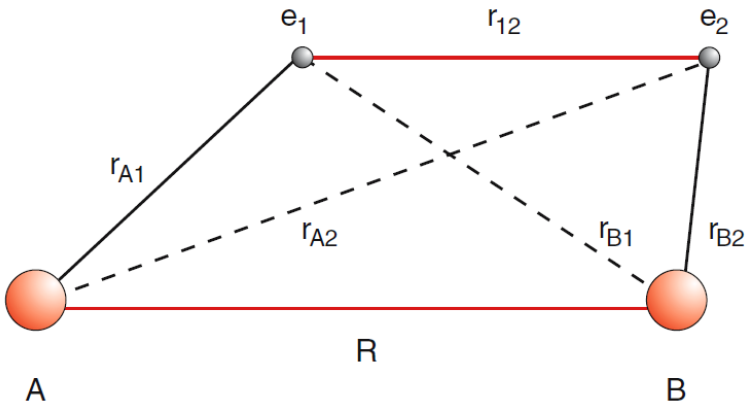

Figure 3.8: The H_2 molecule. From [3]

As a consequence we can no longer separate the wave function into a product of one-dimensional functions, as we could for the H_2^+ ion and no exact analytical solution is possible, even assuming a fixed distance between the nuclei. Furthermore, because two electrons with the same spatial wave functions are non-distinguishable the Pauli principle has to be obeyed. This means the total wave function has to be antisymmetric with respect to an exchange of the two electrons.

The two most commonly used approximations for this molecule are the molecular orbital (MO) approximation and the valence bond theory. They provide good insight into the approximate calculation of electronic states of molecules and their physical motivation. Their differences and possible improvements are discussed in [3].

3.3 Diatomic Molecules

3

Up to now we have discussed the electronic ground state of the molecule H_2^+ and have obtained the potential energy curve $E(R)$ using the LCAO approximation. A similar concept is pursued for larger diatomic molecules with N electrons. The molecular wave function

$$\Phi\left(r_1, \ldots, r_N, R\right) = \Sigma c_n \phi_n \tag{3.21}$$

is always written as a linear combination of many basis functions ϕ_n, which might be atomic orbitals centered at the corresponding nucleus, but also Slater orbitals and Gaussian functions.
The results of the computations, always assuming a fixed distance between the nuclei, give the potential energies $E_k(R)$ for the different electronic states, from which the equilibrium distance R_{ek} can be found.
For a given configuration of the two atoms forming a molecule for decreasing internuclear distance R, many different molecular states are possible.

The different electronic states $E_i(R)$ of diatomic molecules is characterized by the index i, which stands for the various quantum numbers (principal quantum number, angular momentum and spin quantum numbers). The quantum number n characterize the energetic order of the electronic states in atoms, this is not always true for molecules [3].

3.3.1 Electron Configurations and Molecular Ground States

Because the potential, in which the electron is immersed, has no spherical symmetry, the orbital angular momentum l_i is in function of the time and precesses around the symmetry axis of the molecule, which we choose as the z-axis [3]. We define the projection

$$l_z = \lambda\hbar \tag{3.22}$$

by the projection quantum number $\lambda = |m_l|$. The precessing electron produces a magnetic field in the z-direction. For molecules, with one electron and small nuclear charge numbers Z, the interaction between l_i and the electron spin s_i precess independently around the z-axis. The projection of s is called $s_z = \sigma\hbar$, where σ is the spin projection quantum number (Fig. 3.9).

We can characterize the molecular electronic state by the quantum numbers λ, σ and s. For molecules with more than one electron, the coupling order of the different angular momenta depends on the strength of the different interactions.

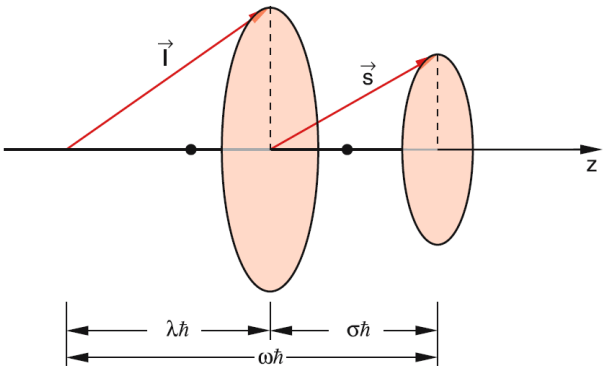

Figure 3.9: Separate precession of s and l around the z-axis for weak spin-orbit coupling. From [3]

If the interaction between different orbital angular momenta l_i is larger than the coupling between l_i and s_i, we are in the so called weak spin orbit coupling situation, the total orbital momentum and the total electron spin are defined as

$$L = \sum l_i \ , \quad S = \sum s_i \tag{3.23}$$

They make an independent precession around the internuclear axis (Fig. 3.10), forming the constant projections

$$L_z = \Lambda \cdot \hbar \quad \text{and} \quad S_z = \Sigma \cdot \hbar \tag{3.24}$$

Electronic states with $\Lambda = 0$ are called Σ-states, with $\Lambda = 1$ are Π-states, with $\Lambda = 2$ are Δ-states, etc.

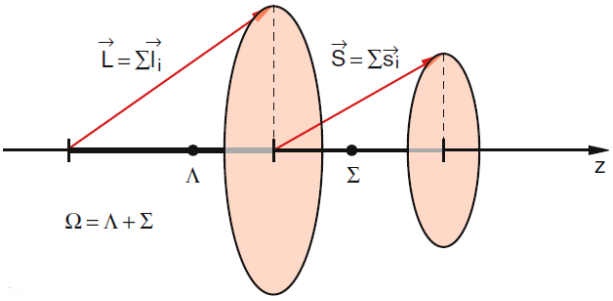

Figure 3.10: Separate precession of $L = \sum l_i$ and $S = \sum s_i$ around the internuclear axis z for small spin-orbit coupling. From [3]

The spins S_A and S_B of the atomic states A and B define the spin

$$S = S_{\mathrm{A}} + S_{\mathrm{B}} \tag{3.25}$$

of the molecular state formed by combination of the two atomic states as the vector sum of the atomic spins. The multiplicity of a molecular state is given by the number $(2S+1)$ of possible projections Σ of S onto the inter-nuclear axis [3].

The molecular orbitals are determined by using wave functions of the type (3.21), they are labeled by the index $n = 1, 2, 3, \ldots$. The orbital angular momentum quantum number l of the electron and its projection quantum number λ, the nomenclature is shown in table 3.3.
The calculations give the following energetic order of the orbitals $\psi(n, l, \lambda)$:

$$1s\sigma, 2s\sigma, 2p\sigma, 2p\pi, 3s\sigma, 3p\sigma, 3p\pi, 3d\sigma, 3d\pi, 3d\delta, 4s\sigma, \ldots,$$

Quantum numbers			
n	l	λ	Term nomenclature
1	0	0	$1s\sigma$
2	0	0	$2s\sigma$
2	1	0	$2p\sigma$
2	1	1	$2p\pi$
3	3	1	$3d\pi$
3	3	2	$3d\delta$

Table 3.3: Quantum numbers and term nomenclature of a molecular electron with principal quantum number n, angular momentum quantum number l, projection quantum number $\lambda = |m_l|$

where the atomic orbital is indicated with the Latin letter and gives the major contribution to the molecular orbital.
The number of electrons in an occupied molecular orbitals are written with an exponent 1 or 2. Orbitals which are left unchanged by the operation of inversion are called gerade, otherwise ungerade. Homo-nuclear diatomic molecules have twice as many orbitals as for hetero-nuclear molecules. This happens because their orbitals can be realized with gerade as well as with ungerade symmetry, i.e., σ_g and σ_u.
Table 3.4 shows the electron configurations, equilibrium inter-nuclear distances R_{e} and binding energies of the ground states for some molecules. where σ_g orbitals are called *bonding orbitals*. They lead to negative bonding energy and are responsible for the creation of molecules. Instead, σ_u orbitals are called *anti-bonding orbitals*. They result in repulsive potential curves that lead to unstable molecules.

Molecule	Configuration	state	R_e/a.u.	E_B/eV
H_2	$(1s\sigma_g)^2 \uparrow\downarrow$	$1\Sigma^+g$	1.398	4.476
He_2^+	$(1s\sigma_g)^2 (1s\sigma_u) \uparrow$	$2\Sigma^+{}_u$	2.041	2.6
He_2	$(1s\sigma_g)^2 (1s\sigma_u)^2 \uparrow\downarrow$	$1\Sigma^+g$	–	0
Li_2	$(1s\sigma_g)^2 (1s\sigma_u)^2 (2s\sigma_g)^2 \uparrow\downarrow$	$1\Sigma^+g$	5.046	1.03
B_2	$(1s\sigma_g)^2 (1s\sigma_u)^2 (2s\sigma_g)^2 (2s\sigma_u)^2 (2p\pi_u)^2 \uparrow\uparrow$	$3\Sigma^-g$	3.005	3.6
LiH	$(1\sigma)^2(2\sigma)^2$	$1\Sigma^+$	3.024	2.52
CH	$(1\sigma)^2(2\sigma)^2(3\sigma)^2 1\pi$	2Π	2.116	3.65
HF	$(1\sigma)^2(2\sigma)^2(3\sigma)^2 1\pi$	$1\Sigma^+$	1.739	6.11

Table 3.4: Electron configurations, equilibrium distances R_e and binding energies of the ground states of some homonuclear and heteronuclear diatomic molecules

In the H_2 molecule, both electrons are in the bonding $1\sigma_g$ orbital. Without the repulsion between the two electrons, the binding energy $E_B\,(R_e)$ of H_2 should be twice as large as for H_2^+. The positive repulsion energy decreases the binding energy slightly. Respect to the H_2^+ molecule, the equilibrium distance R_e is slightly smaller.

The He_2^+ molecule has three electrons; the third electron has to occupy the next higher orbital $1\sigma_u$, which is anti-bonding, as a consequence, its binding energy and equilibrium distance are smaller than that of H_2.
Instead, for the neutral He_2 molecule, which the electronic configuration is shown in the diagram in Fig. 3.11, the binding contribution of the two electrons in the $1\sigma_g$ orbital is nearly cancelled by the anti-bonding energy of the $1\sigma_u$ orbital, therefore, the binding energy of He_2 is extremely small. For temperatures $T > 1\ K$ its ground state is unstable. Finally, the Li_2 molecule has, in addition to the He_2 configuration, two electrons in the bonding $2s\sigma_g$ orbital which contribute to the binding energy. The ground state of Li_2 is therefore stable and has a binding energy of $E_B = -1$ eV [3].

3.3.2 Excited Molecular States

Excited electronic states are formed when one of the electrons occupies an orbital with higher energy; their energies $E_n(R)$ depend on the distance between the nuclei R. We obtain for each of these states, always considering the clamped nuclei approximation, a potential energy curve $E(R)$, which converges for $R \to \infty$ towards the degeneration into a ground state atom and an excited atom.

In general, there are many bound excited states, as can be seen in Fig. 3.12, which shows

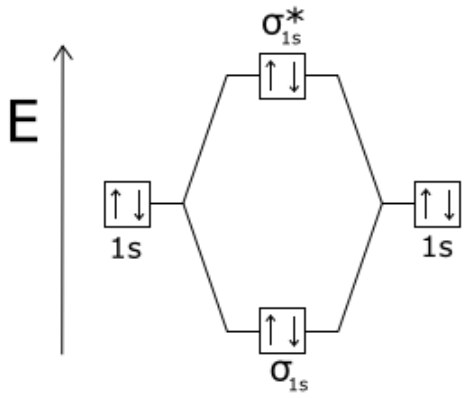

Figure 3.11: Molecular Orbital Diagram for the He_2 molecule. From [17]

potential curves of the Li_2 molecule. When two electrons are excited into higher orbitals, doubly excited states can also be generated. The excited Li_2 molecule dissociates for $R \to \infty$ into two excited Li-atoms. For all these states, the core electrons in the $1s$ atomic shells do not contribute to the molecular binding; they are barely affected by the excitation of one or two electrons from the outer $2s\sigma$ orbital into higher unoccupied orbitals.

The electrons in closed inner shells barely contribute to the formation of molecules. They stay concentrated around their original nucleus when the molecule is formed. At the same time the electrons from the atomic valence shell are rearranged into molecular orbitals. The valence electrons are responsible for the molecular binding energy. Their spatial rearrangement makes the energy to minimize at a certain inter-nuclear distance R_e, the bond length, which is the reason for the binding of the two atoms into a stable molecule [3].

3.4 Basis Sets

In modern computational chemistry, quantum chemical calculations are performed using a finite basis set [19]. These basis sets are used to approximate one-electron functions that are required for the calculation of the ansatzes of electronic structure problem. The choose of the basis set with the right mathematical properties, is made considering the physics of the problem, can remarkably facilitate the evaluation of the Hamiltonian elements and expectation values.
The quality of the basis set can be measured making a comparison with an hypothetical infinite basis set. These are chosen in function of the nature of the problem, were different

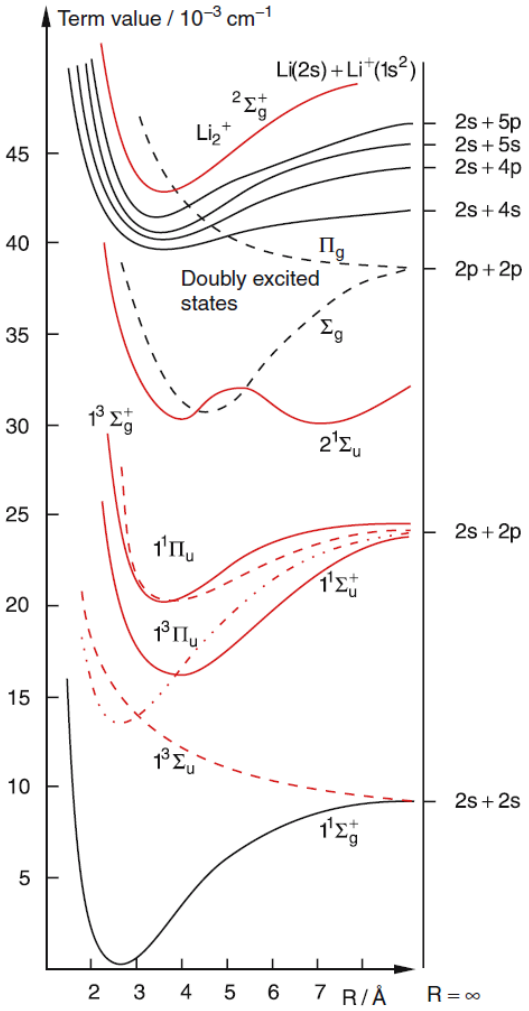

Figure 3.12: Some potential curves of the Li_2 molecule, up to the ionization limit, including some doubly excited states (black dashed curves). From [3]

basis sets offer different levels of compromises. This method provide a systematic way of constructing basis sets for molecules by keeping a dataset of atom orbitals (AOs) for a large part of the elements in the periodic table. The functional form of atomic basis sets is inspired by the solutions of the Schrödinger equation for hydrogen-like atoms, with the general form of the Schrodinger equation (3.3).

The most common basis is the Slater-type orbitals (STOs) basis set. It has the same structure as the orbitals of hydrogenic atoms, with the radial function taking the form

$$R_n^{\mathrm{STO}}(r) = Nr^{n-1}\mathrm{e}^{-\zeta r} \tag{3.26}$$

where N is a normalization constant and ζ is called orbital exponent and controls how fast the density vanishes as a function of the nuclear distance. The minimum and maximum orbital exponents determine how far and how close the resulting wave function can be represented within a basis set comprised of several STOs. With a particular combination of individual basis set functions associated at the orbitals of a selected atom, is possible to construct the AO basis sets, for example, each atom in the second period in the periodic table, a basis set can be constructed using five STOs with different exponents and appropriate spherical harmonics to represent the $1s$, $2s$, $2p_x$, $2p_y$, and $2p_z$ orbitals [19].
An advantage of STOs is that the appropriate behavior of electron density near and far from the nuclei is well predicted. Unfortunately, this basis is restricted to small molecules. This problem the consequence of the lack of analytical solutions for molecular integrals.

To remedy this, the Gaussian-type orbitals (GTOs) basis sets has been developed. It differs in the form of the radial function respect to the (STOs), which adopts a Gaussian form:

$$R_n^{\mathrm{GTO}}(r) = Nr^{n-1}\mathrm{e}^{-\zeta r^2} \tag{3.27}$$

GTOs have convenient analytical properties that allow for an efficient evaluation of the molecular integrals. But, differently from STOs, they fail to describe the exponential tail of the electronic density: this flaw is corrected introducing the contracted GTOs scheme, where a linear combination of GTOs, called primitives, emulate a single STOs. This method is largely used for electronic structure calculations in isolated molecules.
Exist two types of contraction schemes

- Segmented contractions. Different orbitals are represented by different sets of primitive GTOs. An example is the 3-21G, where the core electrons are made of a contraction of 3 GTOs, and the valence electrons are described with contractions of 2 and 1 GTOs.
- General contractions. All the orbitals are expanded using the same set of primitive GTOs, where the combination coefficients of the linear combination are different from each orbital.

In the modern basis sets the schemes segmented and general contraction are mixed.

3.5 Quantum chemistry and quantum computing

Since the advent of classical computers, computation have revolutionized chemistry as a discipline, playing a fundamental role in the development of chemistry. Unfortunately

even supercomputers struggle to model a single molecule in its full complexity and the exponential growth of the dimension of a wave function makes manipulation and storage very inefficient.

However the recent emergence of quantum computation opens a window for solve these problems. With quantum computers is possible to explicit represent the wave function of a system, and differently from a classical computation, the entanglement of the system can be taken in consideration.

Some classical subroutines can be substituted by quantum algorithms that improve the tractability and accuracy of chemical predictions. This brings up the importance of the synergy between quantum information theory and classical quantum chemistry techniques [19].

VQE has been used for the simulation of small molecules as the helium hydride ion [38]. It as been also applied at the beryllium hydride molecule [39]. Using the Variational Quantum Eigensolver and the Google's Sycamore quantum processor, a 12 qubit simulation of a hydrogen chain was demonstrated in 2020 [40].

3.5.1 The second quantized Hamiltonian

The diatomic Hamiltonian is determined by its electronic structure, the kinetic energy of the nuclei is neglected using the clamped nuclei approximation, as justified in Sect. 3.2.1. The first quantized Hamiltonian is described as follows:

$$H_{\text{elec}} = -\sum_i \frac{\nabla^2_{\mathbf{r}_i}}{2} - \sum_{k,j} \frac{Z_k}{|\mathbf{R}_k - \mathbf{r}_j|} + \sum_{k,l>k} \frac{Z_k Z_l}{|\mathbf{R}_k - \mathbf{R}_l|} + \sum_{i,j>i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{3.28}$$

where r_i are the coordinates of the electrons, that depend on the fixed coordinates of the nuclei $\mathbf{R}_k$, and Z_k denote the nuclear charges. An infinite-dimensional Hilbert space is assumed in Eq. (3.28), but for the application of quantum computation, the Hamiltonian must be considered in the second-quantized formulation, in which the system is described approximately using a finite basis, introduced in section 3.4. This basis will be generically denoted with $\phi_i(x)$.

In second quantization states are labeled by the occupation of the orbitals, and the exchange symmetry of the particles is naturally considered through the use of fermionic creation and annihilation operators a_i and $a_i^\dagger$. The second-quantization Hamiltonian is defined as

$$H = h_0 + \sum_{ij} h_{ij} a_i^\dagger a_j + \sum_{ijkl} h_{ijkl} a_i^\dagger a_j^\dagger a_k a_l \tag{3.29}$$

where the constant h_0 represents the nuclear coulombic repulsion. Instead h_{ij} describes the one-electron interaction, and takes in consideration the kinetic energy of each electron and its coulombic attraction towards the two nuclei, it is given by the integral:

$$h_{ij} = \int \mathrm{d}x \phi_i^*(x) \left(\frac{\nabla_r^2}{2} - \sum_k \frac{Z_k}{|\mathbf{R}_k - \mathbf{r}|} \right) \phi_j(x) \tag{3.30}$$

Finally h_{ijkl} represents the two-electron interaction and takes in consideration the Coulomb repulsion between the electrons. It can be written as:

$$h_{ijkl} = \int \mathrm{d}x_1 \, \mathrm{d}x_2 \frac{\phi_i^*(x_1) \, \phi_j^*(x_2) \, \phi_k(x_1) \, \phi_l(x_2)}{|\mathbf{r}_1 - \mathbf{r}_2|} \tag{3.31}$$

Both these coefficients are functions of the inter-nuclear distance, and can be computed using the basis sets introduced in Sect. 3.4.

3.5.2 Mapping to qubits

Once the second quantized Hamiltonian has been obtained, there is a further prerequisite before the Hamiltonian can be implemented in a quantum computer. The creation and annihilation operators a_j and $a_j^\dagger$ have to be mapped to the qubits, where the mapping must preserve the algebraic relations, i.e., the fermionic canonical commutation relations $\left\{ a_j, a_k^\dagger \right\} = \delta_{jk} I$. The principal mapping methods are the Jordan-Wigner transformation and the Bravyi-Kitaev transformation. [19]

Jordan-Wigner transformation

Under the Jordan-Wigner Transform each electronic orbital is described by the annihilation/creation operators that are mapped to qubit operators as follows:

$$a_j \mapsto \tilde{a}_j = \frac{1}{2} (X_j + iY_j) \prod_{i=0}^{j-1} Z_i = (|0\rangle\langle 1|)_j \prod_{i=0}^{j-1} Z_i \tag{3.32}$$

This operator has the following action on a computational basis vector $|z_0, \ldots, z_{N-1}\rangle$.

$$\begin{aligned} \tilde{a}_j \, |z_0 \ldots, z_{j-1}, 1, z_{j+1}, \ldots, z_{N-1}\rangle &= (-1)^{\sum_{i=0}^{j-1} z_i} \, |z_0 \ldots, z_{j-1}, 0, z_{j+1}, \ldots, z_{N-1}\rangle \\ \tilde{a}_j \, |z_0 \ldots, z_{j-1}, 0, z_{j+1}, \ldots, z_{N-1}\rangle &= 0 \end{aligned} \tag{3.33}$$

Observe that $\tilde{a}_p$ acts not only on qubit p but also on qubits $0, \ldots, p-1$. Therefore fermionic operators with low weight can get mapped to qubit operators with high weight, where weight means the number of qubits an operators acts on. These high-weight operators can be problematic, for instance, they may require more gates to simulate and are more expensive to measure on near-term hardware platforms. The annihilation operator on the last mode a_j will map to an operator $\tilde{a}_j$ which acts on all the qubits. [18]

Chapter 4

Diatomic molecule excited states calculation with the QAOA algorithm

In this chapter, we describe the procedure that makes the Quantum Approximation Optimization Algorithm (QAOA) able to calculate the excited states of a generic Hamiltonian. The method is applied to diatomic molecules. It is first tested on the simple case of the Hydrogen molecule, then to the more complicated case of the Lithium Hydride molecule.

We will perform these simulations of the QAOA algorithm on a classical computer. As a consequence, the peculiar errors of quantum computation are not taken into consideration. In a possible implementation of the procedure, experimental errors of the specific platform must be taken into account.

The program is written in the Python programming language v.3.9 [32]. Linear algebra calculations are performed using the *QuTip* library [30]. All the plots are generated by the *Matplotlib* library [31]. The source code is available at `https://github.com/danieletrisciani/qaoa-excited-state`.

4.1 First excited state calculation using the QAOA algorithm

The Quantum Approximate Optimization Algorithm, introduced in Sect. 2.2, is designed to calculate the ground state of the cost Hamiltonian. That can be very limiting if we are interested in excited states of the Hamiltonian. To overcome this problem, to calculate the first excited state, one can add to the original Hamiltonian a penalty term

that gives an extra energy to the ground state. To do so one needs to add to the original Hamiltonian the projector $P_{g.s.} = |\psi_{g.s.}\rangle \langle\psi_{g.s.}|$ over the ground state itself, with a sufficiently large positive coefficient. More specifically, one needs to consider

$$H_{\text{e.s.}} = H + \lambda |\psi_{\text{g.s.}}\rangle \langle\psi_{\text{g.s.}}| \tag{4.1}$$

where $|\psi_{\text{g.s.}}\rangle$ is the ground state calculated using the QAOA algorithm, and λ is a constant that must satisfy the relation:

$$\lambda + \alpha_0 > \alpha_1 \tag{4.2}$$

where α_0 is the ground state energy and α_1 is the first excited state energy. The value of α_1 is unknown. Therefore the calculation of λ is not trivial. We can adopt these two methods:

- The greatest energy α_{max} of the Hamiltonian can be calculated by putting a minus in front of the gradient used in the classical optimizer. We set $\lambda > \alpha_{max}$, and the condition (4.2) is satisfied.
- If the order of magnitude of the difference between the energies is known, we set λ greater than it.

The ground state of $H_{e.s.}$ and its energy are equal to the first excited state of H and its energy. Now the QAOA algorithm can find the excited state energy of the Hamiltonian H, using as cost Hamiltonian the Hamiltonian $H_{e.s.}$.

This procedure can be repeated to calculate every excited state of the Hamiltonian. For instance, to calculate the second excited state, at first, we measure the ground state, and choose a suitable value for λ, we apply the transformation (4.5) that gives us a new cost Hamiltonian for the QAOA that let us calculate the first excited state. Once this state is calculated, it can be used in transformation with a suitable value for λ, the resulting Hamiltonian has in the ground state the second excited state we are looking for. This procedure comes with a problem. Each state has linked to an error; therefore at each iteration of the transformation, more and more errors add up to the result we are looking for.

The QAOA algorithm is performed with the following settings. Values of p range from 1 to 5; as we will see later, $p = 5$ is enough to have a good approximation. The Basin-Hopping method, introduced in Sect. 2.3, performs the classical optimization, where its local optimizer is the BFGS algorithm, explained in Sect. 2.3. The number of Basin-Hopping iterations is 10, this means that the algorithm attempts to find a new local minimum 10 times. The parameter Temperature T is equal to 1. The choice of this optimization algorithm is justified by the nature of the QAOA algorithm, and by the

result of the previous chapter.

Optimization methods are not deterministic, the results of different runs can be very different depending on the randomly chosen initial conditions. To overcome this problem, more than one run is usually performed, and the results are averaged between them. In this work, each measure of the energy and state of the Hamiltonian is repeated 10 times for different randomly generated initial parameters. The average $\alpha_{avg} = \sum_i^K \alpha_i$ and the deviation standard are therefore calculated, the latter is defined as follows

$$\sigma = \sqrt{\frac{1}{K}\sum_{i=i}^{K}(\alpha_i - \alpha_{avg})^2} \tag{4.3}$$

where α_i is the energy or the fidelity for the distance d_i and $K = 10$.

4.2 First excited state calculation for the diatomic molecule Hamiltonian

In this section, we will apply the procedure for calculating the excited states to the Hydrogen and Lithium Hydride Hamiltonian. It will be explained how the second quantization Hamiltonian is built and how the number of qubits to deal with these Hamiltonians can be reduced. Then the ground state for these molecules and some inter-nuclear distances will be calculated using the QAOA algorithm. Once the ground states are calculated, they are used to estimate the first excited states.

4.2.1 The diatomic molecule Hamiltonian in second quantization

As we have seen in Sect. 3.5.1, before a Hamiltonian can be used with the QAOA algorithm it should be transformed in a second quantization form. For a diatomic molecule can be used the second quantized Hamiltonian in Eq. (3.29), for which we have to choose a basis set, introduced in Sect. 3.4 for the estimation of the coefficients h_0, h_{ij} and h_{ijkl} and a transformation for the mapping of the creation and annihilation operators, introduced in Sect. 3.5.2. We will use the Jordan-Wigner transformation.

We will consider only the first 2 molecular orbitals. Thus, the molecule can host up to 2 electrons, with spin $\frac{1}{2}$. Each qubit corresponds to a molecular spin orbital in this way:

- The 1° qubit corresponds to spin ↑ electron in the first molecular orbital.
- The 2° qubit corresponds to spin ↓ electron in the first molecular orbital.

- The 3° qubit corresponds to spin ↑ electron in the second molecular orbital.
- The 4° qubit corresponds to spin ↓ electron in the second molecular orbital.

Using this basis, the Hamiltonian (3.28) becomes a 16x16 matrix of the form:

$$\begin{pmatrix}
0 & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & 1 & 1 & \circ & 1 & \circ & \circ & \circ & 1 & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & 1 & 1 & \circ & 1 & \circ & \circ & \circ & 1 & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & 2 & \circ & 2 & 2 & \circ & \circ & 2 & 2 & \circ & 2 & \circ & \circ & \circ \\
\circ & 1 & 1 & \circ & 1 & \circ & \circ & \circ & 1 & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & 2 & \circ & 2 & 2 & \circ & \circ & 2 & 2 & \circ & 2 & \circ & \circ & \circ \\
\circ & \circ & \circ & 2 & \circ & 2 & 2 & \circ & \circ & 2 & 2 & \circ & 2 & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & 3 & \circ & \circ & \circ & 3 & \circ & 3 & 3 & \circ \\
\circ & 1 & 1 & \circ & 1 & \circ & \circ & \circ & 1 & \circ & \circ & \circ & \circ & \circ & \circ & \circ \\
\circ & \circ & \circ & 2 & \circ & 2 & 2 & \circ & \circ & 2 & 2 & \circ & 2 & \circ & \circ & \circ \\
\circ & \circ & \circ & 2 & \circ & 2 & 2 & \circ & \circ & 2 & 2 & \circ & 2 & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & 3 & \circ & \circ & \circ & 3 & \circ & 3 & 3 & \circ \\
\circ & \circ & \circ & 2 & \circ & 2 & 2 & \circ & \circ & 2 & 2 & \circ & 2 & \circ & \circ & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & 3 & \circ & \circ & \circ & 3 & \circ & 3 & 3 & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & 3 & \circ & \circ & \circ & 3 & \circ & 3 & 3 & \circ \\
\circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & \circ & 4
\end{pmatrix} \tag{4.4}$$

Here, the symbols "$1, 2, 3, 4$" denote coefficients in the subspaces with fixed number of electrons, $N = 1, 2, 3, 4$ respectively, while $\circ$ denotes the elements that are zero since the Hamiltonian conserves the number of particles.

We will consider only states with 2 electrons and opposite spins, in this way less qubits are necessary for treating the Hamiltonian, making the problem more computationally manageable. In the next section it will be explained how this reduction of the number of qubits is performed.

4.2.2 Reduction of the Hydrogen molecule Hamiltonian

We are going to consider the simplest of the diatomic molecules, the Hydrogen molecule, introduced in 3.2.3. We have to calculate the coefficients of the Hamiltonian (3.29) for the Hydrogen molecule for a chosen set of inter-nuclear distances. The coefficients are calculated using the open-source software library *OpenFermion* [28], and the computational chemistry package *Psi4* [29], using the basis set STO-3G, introduced in Sect. 3.4.

The inter-nuclear are chosen in a way to catch the repulsive and attractive behavior of the 2 protons. The bond length for H_2 of 1.4 *a.u.*, taken from the Table (3.4), can help us to estimate them. The distances considered are

$$\text{from } 0.47\ a.u. \text{ to } 4.63\ a.u., \text{ step } 0.09\ a.u. \tag{4.5}$$

In total 45 inter-nuclear distances have been considered. Let us explicitly show the Hamiltonian (4.4) for the Hydrogen molecule for the bond length:

$$H^{(4)} = \begin{pmatrix}
0.706 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.224 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.224 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.439 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.182 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -0.541 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -0.543 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -0.361 & 0 & 0 & -0.182 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.334 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.542 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -0.182 & 0 & 0 & -0.361 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.543 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.334 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.182 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1.116 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.455 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -0.455 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.901
\end{pmatrix}$$

with the corresponding eigenvalues, that can be calculated classically:

$$(-1.13712,\ -0.54278,\ -0.54278,\ -0.54278,\ -0.54171,\ -0.54171,\ -0.45524,\ -0.45524,\ -0.17924,\ 0.2243,\ 0.2243,\ 0.33374,\ 0.33374,\ 0.4598,\ 0.70557,\ 0.90148)\ a.u.$$

To reduce the complexity of the problem, we consider only the subspace with 2 electrons with opposite spins, which is expected to contain the ground state and the first excited state. We can apply the projector $P_2^{(4)}$ at the Hamiltonian. As stated in [1], it is defined as follows

$$P_2^{(4)} = N_\uparrow\,(2 - N_\uparrow)\,N_\downarrow\,(2 - N_\downarrow) \quad \text{and} \quad N_\uparrow = n_1 + n_2, \quad N_\downarrow = n_3 + n_4$$

where $n_j = a_j^\dagger a_j$ is the number operator for the j-th qubit. The projector $P_2^{(4)}$ is applied to the Hamiltonian using this transformation

$$H_N^{(K)} = P_K^{(K)\dagger} H^{(K)} P_N^{(K)}$$

where K is the number of qubits and N is the number of electrons we are considering, in this case $K = 4$ and $N = 2$. The Hamiltonian after the transformation is:

$$H_2^{(4)} = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.43884 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0.18177 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -0.36101 & 0 & 0 & -0.18177 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -0.18177 & 0 & 0 & -0.36101 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.18177 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1.11615 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

Its dimension can be further reduced. Swapping some rows and columns, all the non-zero coefficients can be rearranged in a 4x4 block, and this 4x4 block can be extracted to obtain a 2-qubit Hamiltonian

$$H_2^{(2)} = \begin{pmatrix}
0.43884 & 0.18177 & 0 & 0 \\
0.18177 & -1.11615 & 0 & 0 \\
0 & 0 & -0.36101 & -0.18177 \\
0 & 0 & -0.18177 & -0.36101
\end{pmatrix} \tag{4.6}$$

In [1] arithmetical steps for the extraction of the 4x4 block are illustrated. The reduction of the complexity of the Hamiltonian remarkably speeds up the calculation of its energies for both classical and quantum methods. Therefore the Hamiltonian (4.6) can be decomposed in a two 1-qubit Hamiltonians, making the problem of finding the energies too trivial; this issue will be discussed later.

For the bond length (1.4 *a.u.*), the energies calculated with the diagonalization method are:

$$(-1.13712,\ -0.54278,\ -0.17924,\ 0.4598)\ a.u. \tag{4.7}$$

The form of the matrix Hamiltonian (4.6) does not change if we consider other internuclear distances. This property let us write the Hamiltonian in the general form:

$$H_2^{(2)} = \begin{pmatrix} a & b & 0 & 0 \\ b & c & 0 & 0 \\ 0 & 0 & d & -b \\ 0 & 0 & -b & d \end{pmatrix} \tag{4.8}$$

or in Pauli representation

$$H_2^{(2)} = c_1 \mathbb{1}_4 + c_2 \mathbb{1} \otimes Z + c_2 Z \otimes Z + c_3 Z \otimes X + c_4 Z \otimes \mathbb{1} \tag{4.9}$$

where $a = c_1 + 2c_2 + c_4$, $b = c_3$, $c = c_1 - 2c_2 + c_4$ and $d = c_1 - c_4$.

The orthonormal basis used to write the 2-qubit Hamiltonian has the following meaning:

- $|00\rangle$: spin $\uparrow$ on the second molecular orbital, spin $\downarrow$ on the second molecular orbital
- $|01\rangle$: spin $\downarrow$ on the first molecular orbital, spin $\uparrow$ on the second molecular orbital
- $|10\rangle$: spin $\uparrow$ on the second molecular orbital, spin $\downarrow$ on the first molecular orbital.
- $|11\rangle$: spin $\uparrow$ the first molecular orbital, spin $\downarrow$ on the first molecular orbital.

The energy as a function of the inter-nuclear distance, calculated with the classical diagonalization method, is plotted in Fig. 4.1. The meaning of the states are, in energetic order, illustrated as follow:

- $|\psi_0\rangle = \alpha\,|00\rangle + \beta\,|01\rangle$, where the value $|\alpha|^2$ goes to 0 and $|\beta|^2$ goes to 1 when the inter-nuclear distance tends to 0.
- $|\psi_1\rangle = \frac{1}{\sqrt{2}}\,|11\rangle + |01\rangle$, it means that the probability to have an electron spin $\uparrow$ in the first molecular orbital and another with spin $\downarrow$ in the second molecular orbital is the same if we exchange the electrons.
- $|\psi_2\rangle = \frac{1}{\sqrt{2}}\,|11\rangle - |01\rangle$, it means that the probability to have an electron spin $\uparrow$ in the first molecular orbital and another with spin $\downarrow$ in the second molecular orbital is the same if we exchange the electrons.
- $|\psi_3\rangle = \beta\,|00\rangle + \alpha\,|01\rangle$, where $|\alpha|^2$ goes to 0 and $|\beta|^2$ goes to 1 when the inter-nuclear distance tends to 0.

In the next section the energies of the Hamiltonian (4.1) will be extracted using the QAOA algorithm.

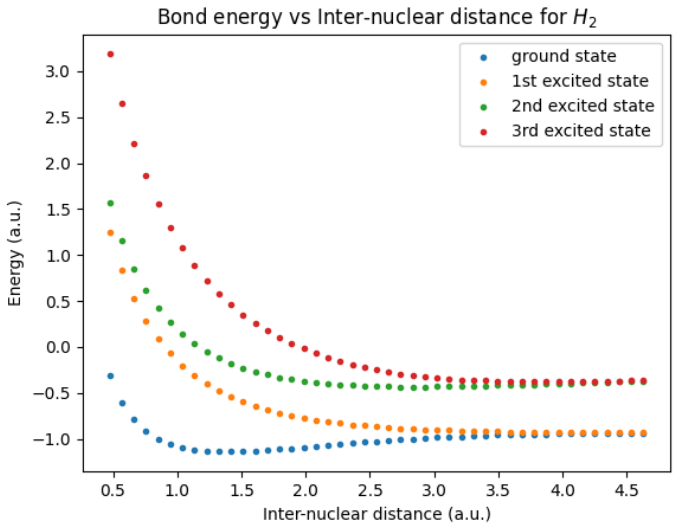

Figure 4.1: H_2 molecule. Energy for all the states of hydrogen molecule as a function of the inter-nuclear distance obtained via exact diagonalization. Only states with 2 electrons with opposite spins in the first 2 molecular orbitals are considered.

4.2.3 First excited state calculation of the H_2 Hamiltonian

The transformation (4.1) requires $|\psi_{\text{g.s.}}\rangle$ to be known. In this section the QAOA algorithm will be used for the calculation of the ground state and then the first excited state of the H_2 Hamiltonian.

The ground state energy of the H_2 Hamiltonian is calculated by the QAOA algorithm for the inter-nuclear distances (4.5). These energies are plotted in Fig. 4.2, for different parameters of the depth parameter $p = 1, ..., 5$.

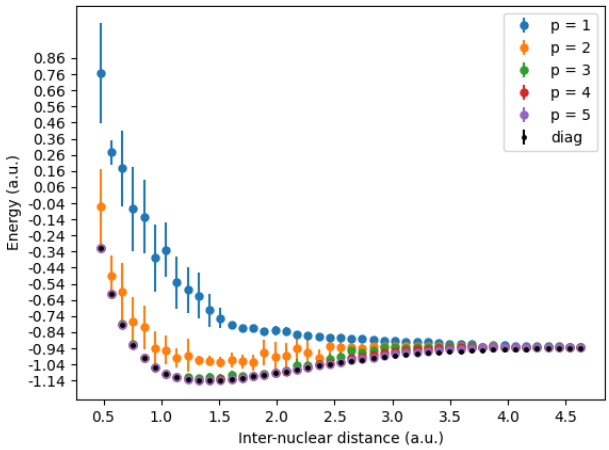

Figure 4.2: H_2 molecule. Ground state energy as a function of the inter-nuclear distance. For values of p from 1 to 5. The calculation is performed by the QAOA algorithm. The energy calculated using the exact diagonalization method are also plotted, to make a comparison. Error bars are calculated by the Eq. (4.3).

As expected, increasing p, we find a better convergence between the energies calculated by the QAOA algorithm and the exact diagonalization. For p values greater than 2, we already have a good convergence. The estimated bond length is 1.41729 $a.u.$ in agreement with that calculated by classical diagonalization and similar to the experimental value in the Table 3.4.

The convergence is estimated by looking at the fidelity $|\langle\psi_{qaoa}|\psi_{qaoa}\rangle|^2$, where $|\psi_{qaoa}\rangle \equiv |\psi(\vec{\beta}_{opt}\vec{\gamma}_{opt})\rangle$ is the QAOA estimated ground state, and $|\psi_{diag}\rangle$ is the classically computed state. It is shown in function of the inter-nuclear distance, and for each value of p, in the plot of Fig. 4.3, instead, Table 4.1 reports the QAOA calculated states and energies for the bond length and for the values for p, together with the exact diagonalization calculated state and energy. The increase of the convergence with p is clearly highlighted in Fig.s 4.4, where the fidelity and the normalized energy are in function of p for a fixed inter-nuclear distance, the bond length. The normalized energy is defined as $\alpha_{norm} = (\alpha - \alpha_{class})/\alpha_{class}$, where α_{class} is the energy calculated classically.

p	$\|\psi_{qaoa}\rangle$	QAOA energy (a.u.)
1	0.07592+0.12544j, 0.79115, 0.17809+0.55896j, 0.08578+0.03265j	-0.7016
2	-0.09116-0.05936j, 0.96828, 0.04522-0.01660j, 0.18484+0.1189j	-1.01403
3	-0.11458, 0.99341, 0, 0	-1.11496
4	-0.11458, 0.99341, 0, 0	-1.13712
5	-0.11458, 0.99341, 0, 0	-1.13712
	$\|\psi_{diag}\rangle$	diag. energy (a.u.)
	-0.11458, 0.99341, 0, 0	-1.13712

Table 4.1: Ground state of the H_2 molecule. The states and energies calculated using the QAOA algorithm are compared with the ones calculated with exact diagonalization, for some values of p, for the bond length distance of 1.42 *a.u.*

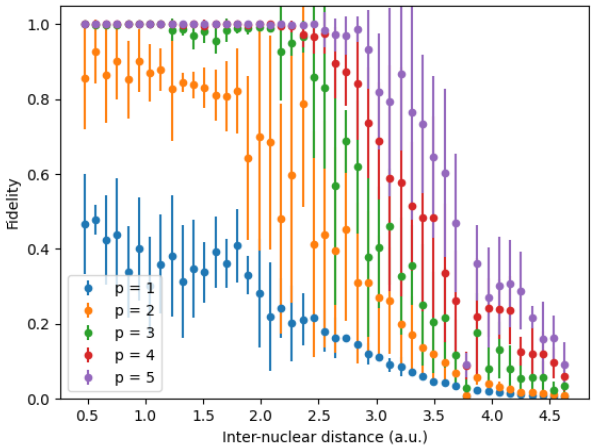

Figure 4.3: H_2 molecule. The fidelity between classically and QAOA calculated ground states in function of the distance for each value of p. With the increasing distance, the classical optimizer can't distinguish between the first excited state and the ground state, resulting in a very low fidelity. Error bars are calculated by the Eq. (4.3).

The plot in Fig. 4.3 clearly shows how the fidelity goes down when the inter-nuclear distance goes to ∞, especially for low values of p. This behavior results from the degeneracy between the ground state and the first excited state energies when the diatomic molecule dissociates in two atoms, as shown in Fig. 4.1. This problem can be partially solved increasing the value of p or the iterations of the basin-hopping algorithm. However, when the molecule dissociates completely, the ground state and the first excited state are indistinguishable.

Now that the ground state is known, the transformation (4.1), that gives us a Hamilto-

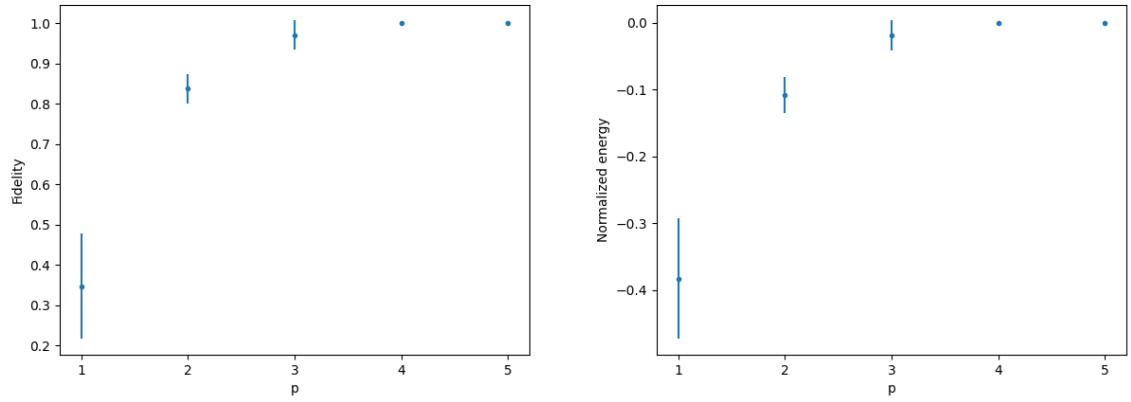

Figure 4.4: Ground state of the H_2 molecule. On the left side, the fidelity $|\langle\psi_{diag}|\psi_{qaoa}\rangle|^2$ for the bond length as a function of p. On the right side, the normalized energy for the bond length as a function of p. Error bars are calculated by the Eq. (4.3).

nian $H_{e.s.}$ which its ground state is the first excited state we are looking for. The order of magnitude of the energies is very low. We have used $\lambda = 10$, it is a valid value because is greater than the difference between the energies of the Hamiltonian, that is usually less than 1 $a.u.$. This new Hamiltonian $H_{e.s.}$ is used as the cost Hamiltonian in the QAOA algorithm. The first excited state energy as a function of the inter-nuclear distance for all the values of p is plotted in Fig. 4.5.

The fidelity for the first excited state is shown as a function of the inter-nuclear distance in the plot given in Fig. 4.6. We observe a good agreement with the values obtained by exact diagonalization only for small inter-nuclear distances. This is because the QAOA calculated ground states for large distances were not well calculated that is the consequence of the degeneracy between the ground state and the first excited state. In Fig.s 4.7 the fidelity and the energy are plotted in function of p, where the values refer to the bond length. In Table 4.2 are reported QAOA calculated excited states and energies for the values of p, considering only the values for the bond length.

From the H_2 Hamiltonian (4.6), it can be deduced that the calculation of its energies can be reduced to a 1-qubit problem (second order equation problem), making it too trivial. A more complex Hamiltonian can be constructed considering another diatomic molecule. All the homogeneous diatomic molecules, with 2 electrons with opposite spin, can be reduced to a 1-qubit problem, as stated in Sect. 4.1. A valid choice for our purpose is the Lithium hydride molecule. In the next section, the QAOA algorithm will be applied to this molecule to calculate the first excited state.

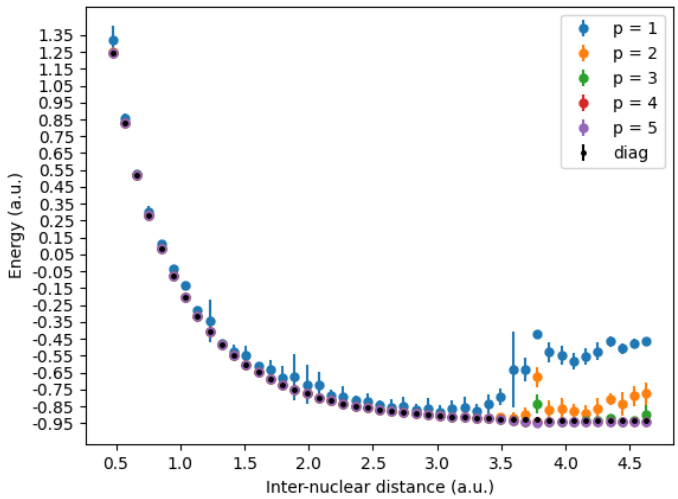

Figure 4.5: H_2 molecule. First excited state energy as a function of the inter-nuclear distance, for values of p from 1 to 5. The calculation are performed by the QAOA algorithm. The energies calculated using the classic diagonalization method are also plotted, to make a comparison. Error bars are calculated by Eq. (4.3).

p	$\|\psi_{qaoa}\rangle$	QAOA energy (a.u.)
1	0.01022-0.03066j, 0.00877+0.00032j, 0.72138, 0.69173-0.00138j	-0.52337
2	0, 0, 0.70711, 0.70711	-0.54273
3	0, 0, 0.70711, 0.70711	-0.54278
4	0, 0, 0.70711, 0.70711	-0.54278
5	0, 0, 0.70711, 0.70711	-0.54278
	$\|\psi_{diag}\rangle$	diag. energy (a.u.)
	0, 0, 0.707, 0.707	-0.543

Table 4.2: For the first excited state of the H_2 molecule. The states and energies calculated using the QAOA algorithm are compared with the ones calculated with exact diagonalization, for some values of p, for the bond length distance of 1.42 $a.u.$

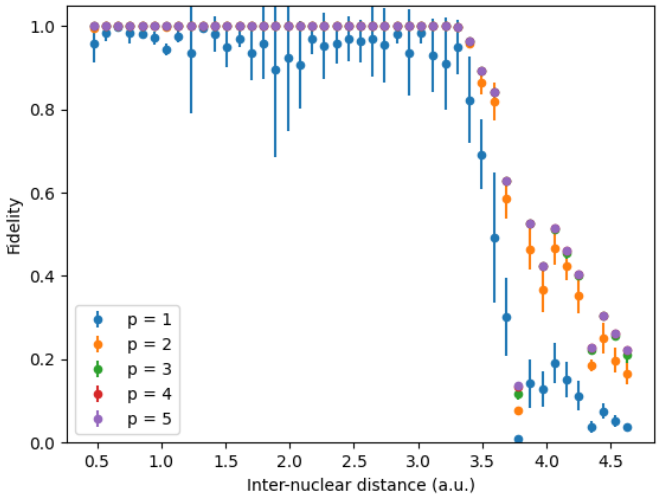

Figure 4.6: H_2 molecule. The fidelity between diagonalization and QAOA calculated excited state in function of the distance, for each value of p. With the increasing distance, the classical optimizer can't distinguish between the first excited state and the ground state, resulting in very low fidelity. Error bars are calculated by the Eq. (4.3).

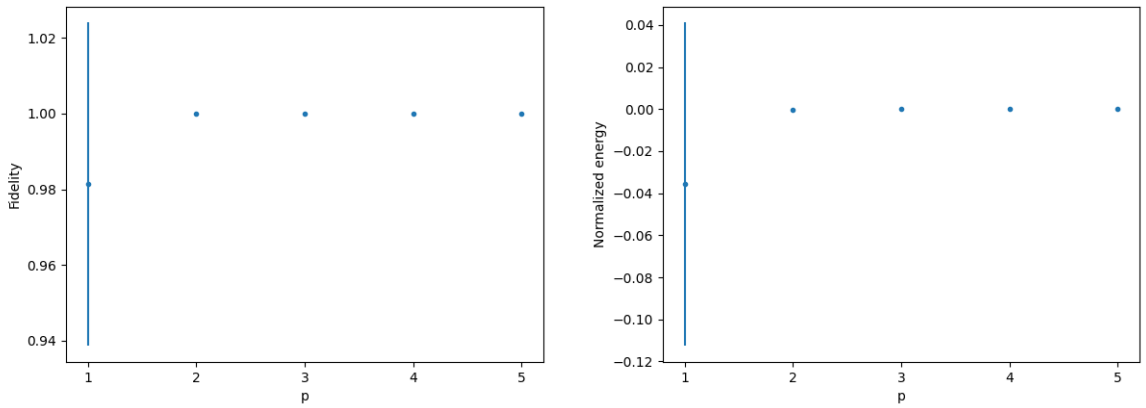

Figure 4.7: Excited states of the H_2 molecule. On the left side, the fidelity $|\langle\psi_{diag}|\psi_{qaoa}\rangle|^2$ for the bond length as a function of p. On the right side, the normalized energy for the bond length as a function of p. Error bars are calculated by the Eq. (4.3).

4.2.4 First excited state calculation of the LiH Hamiltonian

Similarly to the Hydrogen molecule, the STO-3G basis set is used for the construction of the second-quantized Hamiltonian of the LiH molecule. The considered inter-nuclear distances are

$$\text{from } 1.13\ a.u. \text{ to } 6.24\ a.u., \text{ step } 0.09\ a.u. \tag{4.10}$$

for a total of 54 points. The energies as a function of the inter-nuclear distance, calculated with the exact diagonalization method, are plotted in Fig. 4.8.

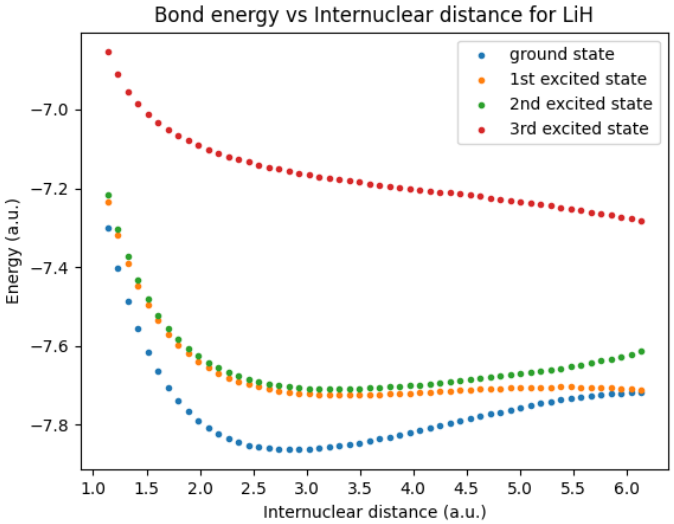

Figure 4.8: LiH molecule. Energies of Lithium hydride as a function of the inter-nuclear distance. Only states with 2 electrons with opposite spins are considered.

For the inter-nuclear distance going to ∞, the ground state and the first excited state become degenerate, as we have seen for the H_2 molecule. Therefore we expect low fidelity for wide distances. From classical calculations is possible to extract the value of the bond length of 2.83459 $a.u.$, similar to the experimental value in the Table 3.4. The reduced Hamiltonian at that distance is

$$H(\ 2.83459\ a.u.) = \begin{pmatrix} -7.16737 & 0.01214 & -0.0531 & 0.0531 \\ 0.01214 & -7.86336 & 0 & 0 \\ -0.0531 & 0 & -7.70529 & -0.01214 \\ 0.0531 & 0 & -0.01214 & -7.70529 \end{pmatrix} \tag{4.11}$$

The form of the matrix Hamiltonian (4.6) does not change if we consider other inter-nuclear distances. This property let us write the Hamiltonian in the general form:

$$H = \begin{pmatrix} a & b & f & -f \\ b & c & 0 & 0 \\ f & 0 & d & -b \\ -f & 0 & -b & d \end{pmatrix} \tag{4.12}$$

or in Pauli representation

$$H = c_1 \mathbb{1}_4 + c_2 \mathbb{1} \otimes Z + c_2 Z \otimes Z + c_3 X \otimes \mathbb{1} - c_3 X \otimes X + c_3 X \otimes Z + c_3 Y \otimes Y + c_4 Z \otimes \mathbb{1} + c_5 Z \otimes X \tag{4.13}$$

where $a = c_1 + 2c_2 + c_4$, $b = c_5$, $c = c_1 - 2c_2 + c_4$, $d = c_1 - c_4$ and $f = 2c_3$.

We can see how the calculation of the energy cannot be decomposed in a 1-qubit problem, as it was for the H_2 molecule Hamiltonian. The structure of states are similar but different from the ones described for the Hydrogen molecule in the end of section 4.2.2.

As stated before, the calculation of the first excited state requires the knowledge of the ground state. The setting of the QAOA algorithm is the same as described in the ending of the section 4.1. The ground state energies, calculated by the QAOA algorithm, as a function of the inter-nuclear distances 4.10 are shown in the plot 4.9.

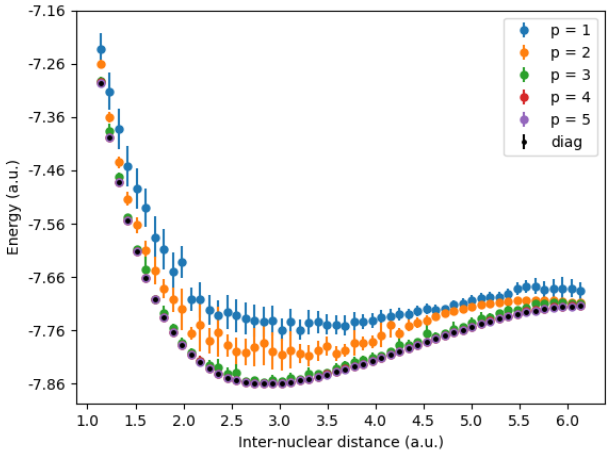

Figure 4.9: Ground state energy, calculated using the QAOA algorithm, as a function of the inter-nuclear distance, for values of p from 1 to 5. The energies calculated using the classic diagonalization method are also plotted to make a comparison. Error bars are calculated by the Eq. (4.3).

For $p > 2$ we have already a good convergence with the classically calculated energies.

The fidelity, plotted in Fig. 4.10 is close to 1 only for $p = 5$. We see there is a greater convergence than in the H_2 molecule. In particular for the inter-nuclear distance going to ∞. For high distance, the fidelity is better than the one obtained for the H_2 molecule. This behaviour results from the longer distance necessary for the total dissociation of the LiH molecule. The fidelity and normalized energy are plotted in Fig.s 4.11 as a function of p, considering only the bond length. As expected, we have better results increasing p. In Table 4.3 are reported QAOA calculated ground states and energies for the values of p, considered the bond length distance.

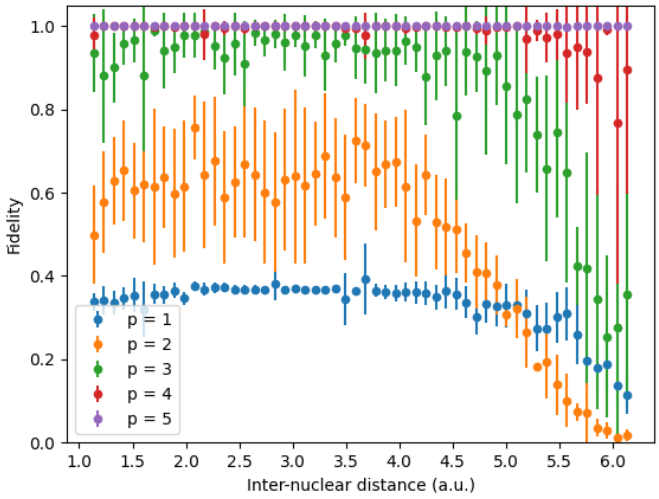

Figure 4.10: LiH molecule. Fidelity between the QAOA and classical calculated ground state, as a function of the inter-nuclear distance, for p from 1 to 5. Error bars are calculated by the Eq. (4.3).

Now that the ground state is known, the transformation (4.1) can be applied on the Lithium Hydride molecule Hamiltonian, allowing the QAOA to approach the first excited state. The first excited state energy as a function of the inter-nuclear distance for all the selected values of p is plotted in Fig. 4.12.
The convergence is high for all the values of $p > 1$, and is better than the one calculated for the ground state. The fidelity for the first excited state as a function of the inter-nuclear distance is plotted in Fig. 4.13. In Table 4.4 are reported QAOA calculated excited states and energies for values of p, considered only for the bond length.
The fidelity and the normalized energy, calculated for the bond length as a function of the values of p, are plotted in the Fig.s 4.14. As expected, the bigger the value of p and better the results. Similarly to the case of the Hydrogen molecule, the excited state fidelity is better than the one calculated for the ground state. We have proven that the

p	$\|\psi_{qaoa}\rangle$	QAOA energy (a.u.)
1	0.17305+0.20231j, 0.70865, 0.01189+0.57262j, -0.2569-0.18142j	-7.74736
2	0.0095-0.08181j, 0.90062, 0.23726-0.21206j, 0.15342+0.23939j	-7.78753
3	-0.01831, 0.9999, -0.00574, 0.00568	-7.86056
4	-0.01831, 0.9999, -0.00569, 0.00571	-7.86357
5	-0.0183, 0.9999, -0.00571, 0.0057	7.86358
	$\|\psi_{diag}\rangle$	diag. energy (a.u.)
	0.0183, -0.9999, 0.0057, -0.005702	-7.86358

Table 4.3: For the ground state of the LiH molecule. The states and energies calculated using the QAOA algorithm are compared with the ones calculated with exact diagonalization, for some values of p, for the bond length distance of 2.83 $a.u.$

p	$\|\psi_{qaoa}\rangle$	QAOA energy (a.u.)
1	0.07097-0.06943j, -0.01204+0.00588j, 0.72895, 0.67403-0.06542j	-7.69685
2	0.01113+0.00487j, 0-0.000821j, 0.70268-0.05236j, 0.70949	-7.71706
3	0, 0, 0.70712, 0.7071	-7.71743
4	0, 0, 0.70712, 0.7071	-7.71743
5	0, 0, 0.70711, 0.70711	-7.71743
	$\|\psi_{diag}\rangle$	diag. energy (a.u.)
	0, 0, 0.70711, 0.70711	-7.71743

Table 4.4: For the first excited state of the LiH molecule. The states and energies calculated using the QAOA algorithm are compared with the ones calculated with exact diagonalization, for some values of p, for the bond length distance of 2.83 $a.u.$

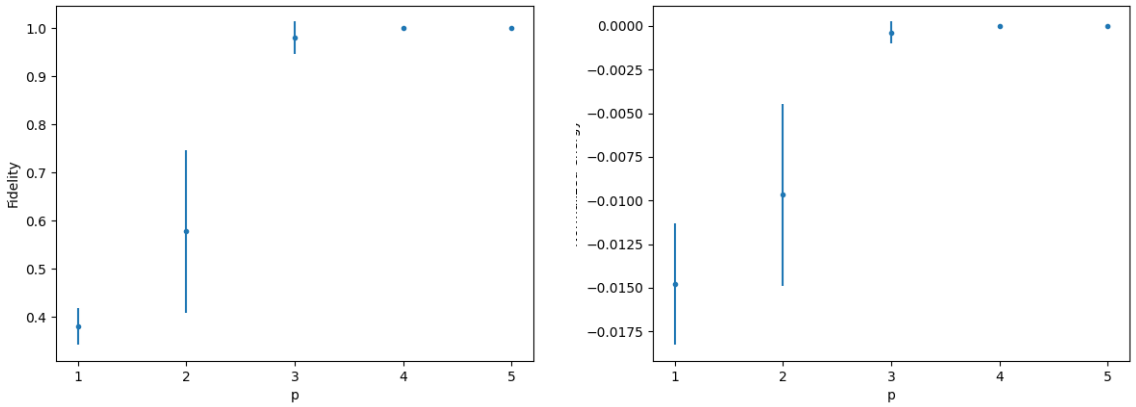

Figure 4.11: Ground states of the LiH molecule. On the left side, the fidelity $|\langle\psi_{diag}|\psi_{qaoa}\rangle|^2$ for the bond length as a function of p. On the right side, the normalized energy for the bond length as a function of p. Error bars are calculated by the Eq. (4.3).

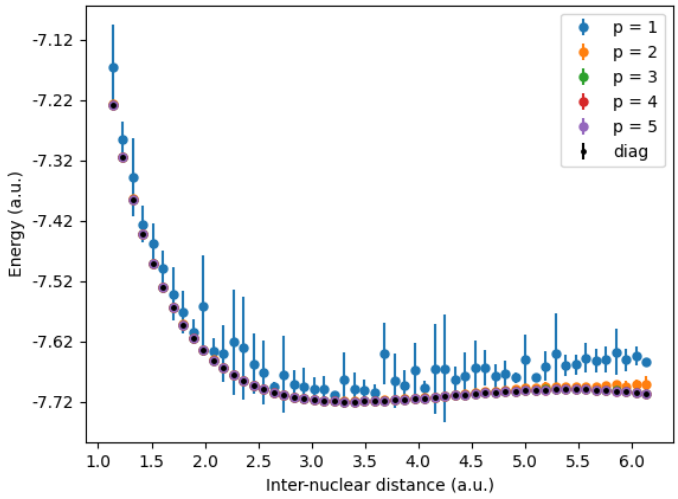

Figure 4.12: LiH molecule. First excited state energy as a function of the inter-nuclear distance, calculated by the QAOA algorithm, for p from 1 to 5. The energies calculated using the classic exact diagonalization method are also plotted to make a comparison. Error bars are calculated by the Eq. (4.3).

QAOA algorithm can calculate the first excited state also for the non-trivial Hamiltonian of the Lithium Hydride molecule.

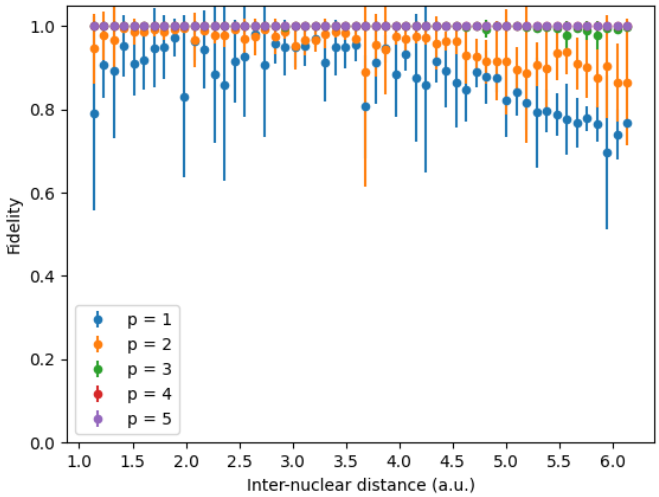

Figure 4.13: *LiH* molecule. Fidelity between the QAOA and classically calculated ground state, as a function of the inter-nuclear distance, for p from 1 to 5. Error bars are calculated by the Eq. (4.3).

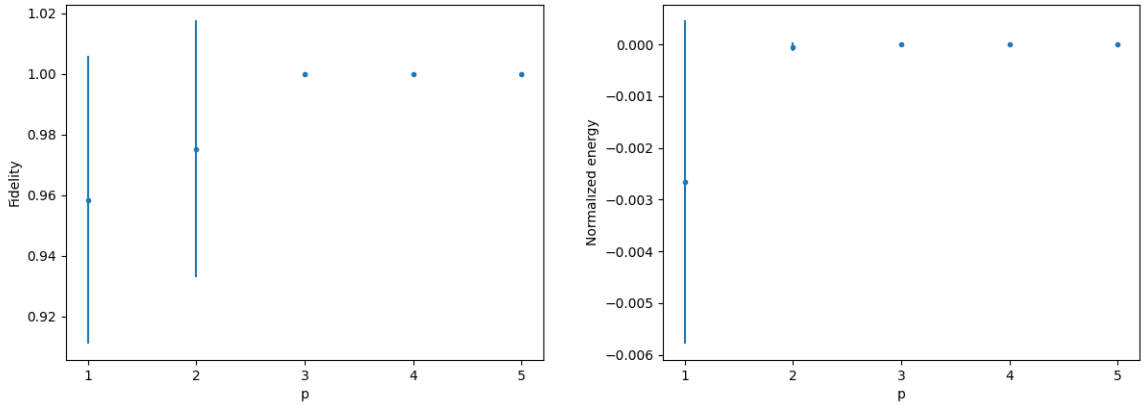

Figure 4.14: Excited states of the *LiH* molecule. On the left side, the fidelity $|\langle\psi_{diag}|\psi_{qaoa}\rangle|^2$ for the bond length as a function of p. On the right side, the normalized energy for the bond length as a function of p. Error bars are calculated by the Eq. (4.3).

Conclusion

The Quantum Approximate Optimization Algorithm was developed initially to solve combinatorial problems, in which the solution is encoded in the ground state of a Hamiltonian. Still, it can be applied to a generic Hamiltonian. The calculation of the excited states using a quantum computer and the QAOA algorithm might be of interest. In this work, we have tested a possible procedure for calculating excited states using the QAOA algorithm. The Basin Hopping method has been used for classical optimization. The values of p considered are from 1 to 5. To measure the first excited state we apply the transformation (4.5) to the Hamiltonian of interest. It put the first excited state and its energy of a Hamiltonian H in the ground state of a Hamiltonian $H_{e.s.}$; in this way, the QAOA algorithm can calculate the first excited state using that latter as the cost Hamiltonian. The transformation requires a parameter λ, which can be each real number greater than the difference between the first excited state and the ground state energies, moreover we have to know the ground state.

This algorithm has been tested using the Hydrogen molecule Hamiltonian, but because of the triviality of calculating its energies, the Lithium Hydride Hamiltonian has been also considered. Only the states with 2 electrons and opposite spins have been considered to reduce the number of qubits necessary. For these molecules, the ground state is calculated by the QAOA algorithm for some distances in the proximity of the bond length. The fidelity between this states and the classically calculated states is also calculated. When the inter-nuclear distance goes to ∞, the first excited state and the ground state become degenerates, and the classical optimizer can't distinguish the two states anymore. The value $p = 5$ is enough to have a good convergence between the states calculated by the QAOA and the diagonalization method, for the distances considered, in the case of the Lithium Hydride. In the case of the Hydrogen, after a certain distance, the fidelity also falls for this value of p; to overcome this problem the value of p or the Basin Hopping iteration should be incremented.

Once the ground state for these distances is known, it can be used to apply the transformation (4.5) to the Hamiltonian and obtain the a new Hamiltonian which the ground state is the first excited state we are looking for. We have chosen $\lambda = 10$; this is an

acceptable value because it is higher than the order of magnitude of the energies. First excited states are calculated using the same values for p, and the same classical optimizer. The ground states used in the transformation are the ones calculated with $p = 5$. In the case of the Lithium Hydride, the first excited state energies and the fidelities are well estimated already with a value of $p = 3$. Differently from the case of the ground state, the fidelity, plotted in Fig. 4.13, doesn't fall for the inter-nuclear distance that goes to ∞; this happens because in this new Hamiltonian, the ground state and the first excited state are not degenerate. Instead, the fidelity for the Hydrogen molecule starts to fall at some distance; this happens because some of the ground states used for the transformation (4.5) were badly estimated because of the degeneration.

In this work, the QAOA is classically simulated. In a real implementation, quantum errors must be taken into consideration.

Bibliography

[1] N. Moll, A. Fuhrer, P. Staar, I. Tavernelli, *Optimizing qubit resources for quantum chemistry simulations in second quantization on a quantum computer*, Journal of Physics A: Mathematical and Theoretical, 49, 29 (2015) URL: https://arxiv.org/abs/1510.04048

[2] E. Farhi, J. Goldstone, and S. Gutmann, *A quantum approximate optimization algorithm* (2014) https://arxiv.org/abs/1411.4028

[3] W. Demtröder, *Atoms, Molecules and Photons* (2010), ISBN : 978-3-642-10297-4 URL: https://link.springer.com/book/10.1007/978-3-642-10298-1

[4] M. A. Nielsen and I. L., *Quantum Computation and Quantum Information*, Cambridge University Press (2000) https://www.cambridge.org/highereducation/books/quantum-computation-and-quantum-information/01E10196D0A682A6AEFFEA52D53BE9AE#overview

[5] D. P. Kingma and J. Ba, *Adam A method for stochastic optimization*, Proceedings of the 3rd International Conference on Learning Representations (ICLR) arXiv:1412.6980 (2017). https://arxiv.org/abs/1412.6980

[6] D. J. Wales, J. P. K. Doye, A. Dullweber, M. P. Hodges, F. Y. Naumkin F. Calvo, J. Hernández-Rojas and T. F. Middleton, *Cambridge Cluster Database*. http://www-wales.ch.cam.ac.uk/CCD.html

[7] D. Wales, J. Doye, *Global Optimization by Basin-Hopping and the Lowest Energy Structures of Lennard-Jones Clusters Containing up to 110 Atoms* J. Phys. Chem. 101, 28, 5111–5116 (1997). https://arxiv.org/abs/cond-mat/9803344

[8] SciPy v1.9.3 Reference Guide (2018-04-20). *scipy.optimize.basinhopping* https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.basinhopping.html

[9] W. Kolos, C.C.J. Rothaan: *Accurate Electronic Wavefunctions for the H2-Molecule, Rev. Mod. Phys. 32, 219* (1960)

[10] W. Kolos and L. Wolniewicz: *Nonadiabatic Theory for Diatomic Molecules and its Application to the Hydrogen Molecule, Rev. Mod. Phys. 35, 473*, (1963)

[11] K. Bharti, A. Cervera-Lierta, T. H. Kyaw, T. Haug, S. Alperin-Lea, A. Anand, M. Degroote, H. Heimonen, J. S. Kottmann, T. Menke, W.-K. Mok, S. Sim, L.-C. Kwek, and A. Aspuru-Guzik, *Noisy intermediate-scale quantum (nisq) algorithms*, Rev. Mod. Phys. 94, 015004 (2022) arXiv:2101.08448

[12] G. E. Crooks, *Performance of the quantum approximate optimization algorithm on the maximum cut problem* (2018) arXiv:1811.08419

[13] L. K. Grover, *"A fast quantum mechanical algorithm for database search"*, Proceedings of the 28th Annual ACM Symposium on the Theory of Computing (1996) arXiv:quant-ph/9605043

[14] R. Fletcher (1987), *Practical Methods of Optimization (2nd ed.)*, New York: John Wiley & Sons, ISBN 978-0-471-91547-8

[15] A. Lam, *BFGS in a Nutshell: An Introduction to Quasi-Newton Methods* Medium (2020) https://towardsdatascience.com/bfgs-in-a-nutshell-an-introduction-to-quasi-newton-methods-21b0e13ee504

[16] S. Patrikar, *Batch, Mini Batch & Stochastic Gradient Descent* Medium (2019) https://towardsdatascience.com/batch-mini-batch-stochastic-gradient-descent-7a62ecba642a

[17] Wikimedia Commons, the free media repository, *He_2 antibonding orbital* https://commons.wikimedia.org/wiki/File:He2_antibonding_orbital.png

[18] Google Quantum AI OpenFermion, *The Jordan-Wigner and Bravyi-Kitaev Transforms* https://quantumai.google/openfermion/tutorial/jordan_wigner_and_bravyi_kitaev_transforms

[19] Y. Cao, J. Romero, J. P. Olson, M. Degroote, P. D. Johnson, M. Kieferová, I. D. Kivlichan, T. Menke, B. Peropadre, N. P. D. Sawaya, S. Sim, L. Veis, and A. Aspuru-Guzik. *Quantum Chemistry in the Age of Quantum Computing* Chemical Reviews 2019 119 (19), 10856-10915 DOI: 10.1021/acs.chemrev.8b00803

[20] D. P. DiVincenzo, IBM *The Physical Implementation of Quantum Computation* Fortschr. Phys., 48: 771-783 (2000). arXiv:quant-ph/0002077

[21] J. Edmonds and R. M. Karp. *Theoretical improvements in algorithmic efficiency for network problems.* J. Assoc. Computing Machinery, 19:248–264, 1922.

[22] P. W. Shor *Polynomial-Time Algorithms for Prime Factorization and Discrete Logarithms on a Quantum Computer* SIAM J.Sci.Statist.Comput. 26, 1484 (1997) arXiv:quant-ph/9508027

[23] S. Ruder *An overview of gradient descent optimization algorithms* (2016) arXiv:1609.04747

[24] J. Duchi, E. Hazan and Y. Singer *Adaptive Subgradient Methods for Online Learning and Stochastic Optimization* Journal of Machine Learning Research 12, 2121-2159 (2011)

[25] E. Duryea, M. Ganger, W. Hu *Exploring Deep Reinforcement Learning with Multi Q-Learning* Intelligent Control and Automation, 7, 129-144. doi: 10.4236/ica.2016.74012.

[26] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh, A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Yuezhen Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M.D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. Jamie Yao, P. Yeh, A. Zalcman, H. Neven and J. M. Martinis. *Quantum supremacy using a programmable superconducting processor*. Nature 574, 505–510 (2019). https://doi.org/10.1038/s41586-019-1666-5

[27] R. F. Stewart, (1 January 1970). *Small Gaussian Expansions of Slater-Type Orbitals*. The Journal of Chemical Physics. 52 (1): 431–438. doi:10.1063/1.1672702.

[28] J. R. McClean, K. J. Sung, I. D. Kivlichan, Y. Cao, C. Dai, E. Schuyler Fried, C. Gidney, B. Gimby, P. Gokhale, T. Häner, T. Hardikar, V. Havlíček, O. Higgott, C. Huang, J. Izaac, Z. Jiang, X. Liu, S. McArdle, M. Neeley, T. O'Brien, B. O'Gorman, I. Ozfidan, M. D. Radin, J. Romero, N. Rubin, N. P. D. Sawaya, K. Setia, S. Sim, D. S. Steiger, M. Steudtner, Q. Sun, W. Sun, D. Wang, F. Zhang and R. Babbush (2017) *OpenFermion: The Electronic Structure Package for Quantum Computers*. Quantum Sci. Technol. 5 034014 (2020) arXiv:1609.04747

[29] R. M. Parrish, L. A. Burns, D. G. A. Smith, A. C. Simmonett, A. E. DePrince III, E. G. Hohenstein, U. Bozkaya, A. Yu. Sokolov, R. Di Remigio, R. M. Richard, J.

F. Gonthier, A. M. James, H. R. McAlexander, A. Kumar, M. Saitow, X. Wang, B. P. Pritchard, P. Verma, H. F. Schaefer III, K. Patkowski, R. A. King, E. F. Valeev, F. A. Evangelista, J. M. Turney, T. D. Crawford, and C. D. Sherrill. *Psi4 1.1: An Open-Source Electronic Structure Program Emphasizing Automation, Advanced Libraries, and Interoperability*, J. Chem. Theory Comput., 13(7) 3185–3197 (2017) (doi: 10.1021/acs.jctc.7b00174).

[30] J. R. Johansson, P. D. Nation, and F. Nori: *QuTiP: An open-source Python framework for the dynamics of open quantum systems.*, Comp. Phys. Comm. 183, 1760–1772 (2012) [doi: 10.1016/j.cpc.2012.02.021].

[31] J. D. Hunter, *Matplotlib: A 2D Graphics Environment*, Computing in Science & Engineering, vol. 9, no. 3, pp. 90-95, 2007. [doi: 10.5281/zenodo.7637593]

[32] Python Software Foundation. Python Language Reference, version 3.9 Available at http://www.python.org

[33] P. Kok, W.J. Munro, K. Nemoto, T.C. Ralph, J. P. Dowling and G.J. Milburn *Review article: Linear optical quantum computing* Rev. Mod. Phys. 79, 135 (2007). arXiv:quant-ph/0512071

[34] L. M. K. Vandersypen and I. L. Chuang *NMR techniques for quantum control and computation* Rev. Mod. Phys. 76, 1037 (2005) arXiv:quant-ph/0404064

[35] H. Huang, D. Wu, D. Fan and X. Zhu *Superconducting Quantum Computing: A Review* Science China Information Sciences 63, 180501 (2020) arXiv:2006.10433

[36] D. Aharonov and M. Ben-Or *Fault-Tolerant Quantum Computation with Constant Error Rate*. SIAM Journal on Computing. 38, 4, 10.1137/S0097539799359385 (2018) arXiv:quant-ph/9906129.

[37] A. M. Dalzell, A. W. Harrow, D. Enshan Koh, R. L. La Placa. *How many qubits are needed for quantum computational supremacy?*. Quantum 4, 264 (2020) arXiv:1805.05224

[38] A. Peruzzo, J. McClean, P. Shadbolt, M. Yung, X. Zhou, P. J. Love, A. Aspuru-Guzik and J. L. O'Brien *A variational eigenvalue solver on a quantum processor* Nature Communications, 5:4213, (2014) arXiv:1304.3061

[39] A. Kandala, A. Mezzacapo, K. Temme, M. Takita, M. Brink, J. M. Chow and J. M. Gambetta *Hardware-efficient Variational Quantum Eigensolver for Small Molecules and Quantum Magnets* Nature 549, 242 (2017) arXiv:1304.3061

[40] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, S. Boixo, M. Broughton, B. B. Buckley, D. A. Buell, B. Burkett, N. Bushnell, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, S. Demura, A. Dunsworth, E. Farhi, A. Fowler, B. Foxen, C. Gidney, M. Giustina, R. Graff, S. Habegger, M. P. Harrigan, A. Ho, S. Hong, T. Huang, W. J. Huggins, L. Ioffe, S. V. Isakov, E. Jeffrey, Z. Jiang, C. Jones, D. Kafri, K. Kechedzhi, J. Kelly, S. Kim, P. V. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, P. Laptev, M. Lindmark, E. Lucero, O. Martin, J. M. Martinis, J. R. McClean, M. McEwen, A. Megrant, X. Mi, M. Mohseni, W. Mruczkiewicz, J. Mutus, O. Naaman, M. Neeley, C. Neill, H. Neven, M. Yuezhen Niu, T. E. O'Brien, E. Ostby, A. Petukhov, H. Putterman, C. Quintana, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, D. Strain, K. J. Sung, M. Szalay, T. Y. Takeshita, A. Vainsencher, T. White, N. Wiebe, Z. J. Yao, P. Yeh, and A. Zalcman *Hartree-Fock on a superconducting qubit quantum computer*. Science 369 (6507), 1084-1089, 2020.

Acknowledgments

I would like to thank my supervisor Prof. Elisa Ercolessi and my co-supervisor Dr. Claudio Massimiliano Sanavio for their dedicated support through my thesis work. I would also like to thank Dr. Davide Vodola for giving me the opportunity to work on the topic of quantum computation.

Finally, I want to thank my family and my friends for the support, Roberto for hosting me in Bologna when I was looking for an apartment, Alessandro e Pietro for the chats during lunch breaks and my flatmates for the good times we spent together.