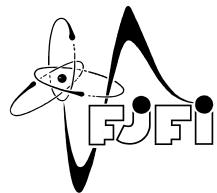


CZECH TECHNICAL UNIVERSITY IN PRAGUE
Faculty of Nuclear Sciences and Physical Engineering



Využití kvantových počítačů v kvantové chemii

Applications of quantum computers in quantum chemistry

Research project

Author:	Ludvík Cigna
Supervisor:	doc. Ing. Martin Štefaňák, Ph.D.
Consultant:	RNDR. Ing. Michal Jex, Ph.D.
Academic year:	2021/2022

- Work assignment -

Acknowledgment:

I would like to thank especially my supervisor doc. Ing. Martin Štefaňák, Ph.D. for his diligence, willingness, friendliness and professional and human support during the guidance my research project and I would like to thank my consultant RNDR Ing. Michal Jex, Ph.D. for the opportunity of professional consultation.

Affidavit:

I declare that I have prepared this work independently and I have listed all the literature used.

In Prague, June 7, 2022

Ludvík Cigna

Title:

Applications of quantum computers in quantum chemistry

Author: Ludvík Cigna

Field of study: Quantum Technologies

Type of thesis: Research project

Supervisor: doc. Ing. Martin Štefaňák, Ph.D., Department of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague

Consultant: RNDR. Ing. Michal Jex, Ph.D., Department of Physics, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University in Prague

Abstract: Quantum computing has received a lot of attention recently, and specifically its application in quantum chemistry looks very promising. This research project summarizes the most important methods used in calculating the electronic structure of a chemical system on quantum computers. This includes the formulation of the problem in first and second quantization and the use of classical computational methods of quantum chemistry, starting with the Born-Oppenheimer approximation. In particular, attention is paid to the Hartree-Fock method, from which more advanced approaches follow. Next, methods for mapping fermions to qubits are discussed and extra space is given to the Bravyi-Kitaev transform. This is then exploited in two quantum algorithms suitable for determining the ground state. The first is an algorithm for phase estimation. Within this, two methods for simulating the Hamiltonian are described, namely the so-called Trotterization and qubitization. The second is a variational algorithm for energy estimation suitable already for NISQ computers. Finally, the derived methods are illustrated in the practical section with the simple example of molecular hydrogen H₂.

Key words: Hartree-Fock method, quantum computing, quantum chemistry, quantum phase estimation algorithm, variational quantum eigensolver

Název práce:

Využití kvantových počítačů v kvantové chemii

Autor: Ludvík Cigna

Abstrakt: Kvantovým počítačům se v poslední době dostává velké pozornosti a konkrétně jejich využití v kvantové chemii se jeví velmi slibně. V této práci jsou shrnutы nejdůležitější metody používané při výpočtu elektronové struktury chemického systému na kvantových počítačích. To zahrnuje formulaci problému v prvním a druhém kvantování a využití klasických výpočetních metod kvantové chemie, počínaje Born-Oppenheimerovou aproximací. Zejména pak je věnována pozornost Hartree-Fockově metodě, z níž plynou pokročilejší přístupy. Dále jsou diskutovány způsoby mapování fermionů na qubity, speciální prostor je udělen Bravyi-Kitaevově transformaci. Toho je pak využito ve dvou kvantových algoritmech vhodných k určení základního stavu. Prvním je algoritmus pro odhad fáze. V rámci toho jsou popsány dvě metody simulace Hamiltoniánu, konkrétně tzv. Trotterizace a qubitizace. Druhým je variační algoritmus pro odhad energie vhodný již pro NISQ počítače. Nakonec jsou odvozené metody ilustrovány v praktické části na příkladu jednoduché molekuly vodíku H₂.

Klíčová slova: algoritmus pro odhad fáze, Hartree-Fockova metoda, kvantová chemie, kvantové počítání, variační algoritmus pro odhad energie

Contents

Introduction	6
I Theoretical part	8
I.1 Problem statement	8
I.2 Fermionic nature, formulation in first and second quantization	9
I.3 Born-Oppenheimer approximation	12
I.4 Hartree-Fock method	14
I.4.1 Variational principle	14
I.4.2 Hartree-Fock equations	15
I.4.3 Numerical calculation, basis sets	18
I.5 Post-Hartree-Fock methods, typical calculation	20
I.6 Mapping to qubits	21
I.6.1 Mapping in the first quantization	22
I.6.2 Mapping in second quantization	23
I.7 Determining the electronic structure using the phase estimation algorithm	27
I.7.1 Initial state preparation	28
I.7.2 Hamiltonian simulation	29
I.8 Ground state determination using VQE	34
I.8.1 Choice of ansatz for state preparation	35
I.8.2 Hardware Efficient Ansatze	39
I.8.3 Energy measurement	40
I.8.4 Classical optimization	43
II Practical part	46
II.1 H ₂ - mapping and reduction of the number of qubits	46
II.1.1 STO-3G	46
II.1.2 6-31G	48
II.1.3 cc-PVDZ	48
II.2 Explicit UCCSD ansatz for STO-3G	49
II.3 PES sampling under various choices of basis sets and mappings	50
II.4 Comparison of optimization techniques, ansatze and mappings	55
Conclusion	60

Introduction

A quantum computer is a designation for a very precisely controlled quantum system from which useful results can be obtained through a well-chosen set of operations and subsequent measurement. It is analogous to classical computers, with the difference that the basic computational unit - the quantum bit called *qubit* - is subject to the laws of quantum mechanics. As such, a quantum computer can be physically realized using a wide range of platforms. Among the most widespread are superconducting circuits [20], ion traps [12], topological qubits [33], or nuclear magnetic resonance [43].

One of the first to see the advantages of such devices was Richard Feynman, when he stated: "If you want to make a simulation of nature, you'd better make it quantum mechanical, and by golly it's a wonderful problem, because it doesn't look so easy." He was referring to the problem of simulating quantum systems on classical computers, where complexity generally grows exponentially with the size of the system studied. The idea that the simulation of quantum systems would be best done using other quantum systems persists today and seems like the most obvious and closest use of quantum computers.

This work focuses on the simulation of quantum chemistry. The central task here is to determine the ground state and its energy. This task is often called the problem of electronic structure. For simulation on classical computers, a huge number of computational methods have been invented, which improved along with the development of increasingly powerful classical computers. The basis of most *ab initio* methods is the Hartree - Fock method (see chapter I.4), on which more accurate models are subsequently built. Among *first principle* methods, *Density Functional Theory* is arguably the most important, which is used in most scientific publications in this area today. However, all these methods encounter the problem of complexity when realistically describing more extensive systems. Due to the natural quantum behavior of quantum computers, there is a general consensus that they will be able to overcome this problem and provide results leading to a deeper understanding of these systems. The application of these findings can then be in chemistry as well as in medicine, biology, or in the design of new materials.

The first possibility to obtain the ground state and its energy is by using the phase estimation algorithm [42]. In principle, this algorithm can also obtain higher energy states. However, the required depth of circuits implementing this algorithm and the resulting long coherence times are problematic. Therefore, to implement this algorithm, a device capable of error correction is needed, which, even by optimistic estimates, will not be available earlier than in the next decade. There is, therefore, an effort to devise algorithms that can provide useful results even on current NISQ devices. The *Variational Quantum Eigensolver* (VQE) appears most promising. Here, deep circuits are replaced by repeated state preparation and measurement. VQE has been implemented by many research teams on many platforms [47], [56], [30].

A number of review articles on quantum computational chemistry have been written, among which this work could be classified. Among the most recent and most important for this work are [15], [57], [37]. While most of these works aim to cover as many known methods and approaches in quantum chemistry as possible and are targeted more at readers who already have some experience with classical computational chemistry methods, this work focuses on a detailed description of the most commonly

used methods in a way understandable to readers without a background in quantum chemistry. Thus, there is an effort instead of describing a wide range of approaches to outline one direction of computation detailed from the physical problem to obtaining results from the final program.

This work is divided into two large parts - theoretical (I) and practical (II). In the theoretical part, section I.1 will introduce the physical problem followed by a summary of fermion behavior and the formulation of the problem in first and second quantization in section I.2. Then, two building blocks of most calculations in quantum chemistry - the Born-Oppenheimer approximation in section I.3 and the Hartree-Fock method in I.4 - will be derived. Advanced methods will be outlined in I.5. Subsequently, the mapping of the physical problem to qubits will be discussed in detail in section I.6. Afterward, the discussion will split according to the method of solving the electronic structure problem using the phase estimation algorithm in I.7 and using VQE in I.8. In the practical part, the derived methods will be demonstrated on the example of molecular hydrogen H₂. Mapping for various basis sets will be demonstrated in section II.1, followed by an explicit demonstration of the chosen ansatz in II.2. Finally, using the quantum programming package Qiskit [2], the potential energy surface calculation will be performed in section II.3, and then different choices of optimization techniques, ansatzes, and mappings will be compared in II.4.

Chapter I

Theoretical part

I.1 Problem statement

The fundamental goal of quantum chemistry is to determine the electronic structure of various molecules. This requires solving the non-relativistic timeless Schrödinger equation for a given molecule

$$H_{mol} |\Psi_n\rangle = E_n |\Psi_n\rangle. \quad (\text{I.1})$$

Here, H_{mol} is the total Hamiltonian of the system, $|\Psi_n\rangle$ the many-particle wave function, and E_n the energy eigenvalue. It is appropriate to mention that relativistic effects cannot always be neglected in quantum chemistry, especially when it comes to heavy elements [49]. This is therefore the first approximation that is considered in this work. The Hamiltonian of a molecule consisting of K nuclei and N electrons is given as

$$\begin{aligned} H_{mol} &= - \sum_i \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 - \sum_{i,I} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \frac{1}{2} \sum_{I,J,I \neq J} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I Z_J}{|\mathbf{R}_I - \mathbf{R}_J|} \\ &= T_e + T_n + V_{en} + V_{ee} + V_{nn}, \end{aligned} \quad (\text{I.2})$$

where M_I is the mass, \mathbf{R}_I the position, and Z_I the number of protons of the I -th nucleus, and \mathbf{r}_i the position of the i -th electron. The first two sums are then the kinetic energies of electrons and nuclei, respectively, and the remaining three sums represent the Coulomb interaction between electrons and nuclei, electrons among themselves, and nuclei among themselves.

An exact solution to such a problem, due to the exponential growth of the wave function's size with the number of particles, goes beyond our computational possibilities [31]. This is the first hint that quantum computers might have an advantage over classical ones. However, before a detailed description of quantum algorithms, the most important approximations that helped classical computers achieve very good results in calculating the electronic structure of many molecules will be discussed. Before we proceed to a detailed discussion of these methods, it should be clarified what exactly we gain by calculating the energy spectrum and especially by calculating the ground state.

The first simplification is the so-called Born-Oppenheimer approximation, which is described in detail in the following chapter. Here, let us just mention that by separating the motion of electrons and nuclei, the original problem I.1 is reduced to finding the eigenstates and eigenvalues of the electronic Hamiltonian

$$H_e = T_e + V_{en} + V_{ee} \quad (\text{I.3})$$

parameterized by the coordinates of the nuclei (for example, their distances or angles between them). In this way, we can reconstruct the so-called potential energy surface (PES) between reactants and products. It allows not only to find the configuration (shape) of the molecule with the lowest energy but also to determine the reaction mechanism (*reaction pathway*). The saddle point on the PES here represents the transition state. Using the transition state theory [22], [21], [64], we are then able to calculate the reaction rate as

$$\text{Reaction rate} \propto \exp\left\{\frac{\Delta E}{k_B T}\right\}, \quad (\text{I.4})$$

where T is the temperature of the system and ΔE is the difference in energy between reactants and products. This is also related to the concept of *chemical accuracy*, which refers to the degree of accuracy in determining the energy needed for the predictions made by computational chemistry to be realistic. It is defined as $1.6 \cdot 10^{-3}$ Hartree, where 1Hartree = 27.211eV is the unit of energy used in quantum chemistry.

I.2 Fermionic nature, formulation in first and second quantization

Our object of interest is electrons. They fall under fermions (obeying Fermi-Dirac statistics) and must comply with the Pauli exclusion principle. This has a fundamental impact on the properties of their wave function. Specifically, it must be antisymmetric with respect to the exchange of two electrons. This fact can be captured in two different formulations - so-called first and second quantization. In the first case, the antisymmetry is maintained directly in the form of the wave function, while in the second, the required statistics are ensured by operators acting on the wave function. It should also be mentioned that in both formulations, computational methods can be divided into so-called lattice-based and based on a set of basis functions. However, this work will consider only the latter of the mentioned approaches.

In first quantization, the question is which particle is in which state. This question is essentially non-physical due to the indistinguishability of particles and introduces redundancy into the description. Therefore, the formulation in second quantization is almost exclusively used in today's classical computational methods. However, in this work, the description in first quantization will be utilized not only due to its explicitness but also in view of its application in quantum algorithms. The antisymmetry of the many-electron wave function can be expressed using so-called Slater's determinant

$$\psi(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1}) \propto \begin{vmatrix} \phi_0(\mathbf{x}_0) & \phi_1(\mathbf{x}_0) & \dots & \phi_{N-1}(\mathbf{x}_0) \\ \phi_0(\mathbf{x}_1) & \phi_1(\mathbf{x}_1) & \dots & \phi_{N-1}(\mathbf{x}_1) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \phi_0(\mathbf{x}_{N-1}) & \phi_1(\mathbf{x}_{N-1}) & \dots & \phi_{N-1}(\mathbf{x}_{N-1}) \end{vmatrix}, \quad (\text{I.5})$$

where $\{\phi_j(\mathbf{x}_i)\}$ is a set of one-electron wave functions (spin-orbitals) in quantum chemistry called molecular orbitals (MOs) and $\mathbf{x}_i = (\mathbf{r}_i, \sigma_i)$ indicates the position and spin variable of the particle. Thus

$$\phi_j(\mathbf{x}_i) = \psi_j(\mathbf{r}_i) \xi_j(\sigma_i) = \begin{cases} \psi_j(\mathbf{r}_i) \alpha(\sigma_i) \\ \psi_j(\mathbf{r}_i) \beta(\sigma_i) \end{cases}, \quad (\text{I.6})$$

where the spin part is described by two orthonormal functions of the spin variable (integrating over them yields the Kronecker delta). Clearly, a set of $2M$ MO $\phi_j(\mathbf{x}_i)$ can be formed using a set of M spatial orbitals $\psi_j(\mathbf{r}_i)$.

The fulfillment of the required properties is obvious - exchanging electrons corresponds to swapping rows, which results in a change of sign, and moreover, if two states were the same, then the determinant would have two identical columns and would therefore be zero. It is common to consider more MOs than there are electrons, so that we can, for example, include excited states in the description. However, this cannot be well captured by writing with Slater's determinant in first quantization. However, realizing that these Slater determinants are eigenstates of the Hamiltonian of the given problem and that it is of course a Hermitian operator, these determinants will surely form a basis in the Hilbert space of this problem. At this point, it suffices to say that the solution can be written as their linear combination, and the appropriate notation will be shown shortly.

The formulation in second quantization stems from the first quantization and was developed primarily for the quantization of fields in quantum field theory. Besides providing a more compact notation, its main advantage is the ability to work with a variable number of particles. Again, consider N electrons and now M molecular orbitals ϕ_j , $M \geq N$ (for completeness, one could consider M to be infinite, but the goal is representation on a computer, so a finite number of MOs is considered). On the space of N indistinguishable particles \mathcal{H} , an orthogonal projector onto the subspace of antisymmetric states \mathcal{H}_A , denoted A , can be introduced. It can be easily shown that if we define it as

$$A = \frac{1}{N!} \sum_{\pi} \sigma_{\pi} P_{\pi}, \quad (\text{I.7})$$

where P_{π} is the permutation operator π and σ_{π} is its sign, then

$$A |\phi_{N-1} \dots \phi_0\rangle = \frac{1}{N!} \sum_{\pi} \sigma_{\pi} P_{\pi} |\phi_{N-1} \dots \phi_0\rangle \in \mathcal{H}_A. \quad (\text{I.8})$$

The Slater determinant can already be seen in this, and these states form the ON basis in \mathcal{H}_A . Obviously, the correct normalization here is $\sqrt{N!}$. Let us therefore denote the normalized antisymmetrized state

$$|\phi_{N-1} \dots \phi_0; A\rangle = \sqrt{N!} A |\phi_{N-1} \dots \phi_0\rangle. \quad (\text{I.9})$$

Finally, the occupation numbers $\{n_j\}$, $j \in \{0, \dots, M-1\}$ are introduced, indicating the number of particles in the state ϕ_j (1 if it is occupied and 0 vice versa). They have to satisfy

$$\sum_{j=0}^{M-1} n_j = N. \quad (\text{I.10})$$

From the Pauli exclusion principle $n_j \in \{0, 1\}$ will hold for electrons. The state of N electrons in M molecular orbitals will then be

$$|n_{M-1}, \dots, n_0; A\rangle = \sqrt{N!} A (|\phi_{M-1}\rangle^{\otimes n_{M-1}} \otimes \dots \otimes |\phi_0\rangle^{\otimes n_0}) \equiv |\{n\}\rangle \quad (\text{I.11})$$

and this will henceforth be called Slater's determinant in second quantization. Since such Slater determinants are eigenstates of the Hermitian operator, they form a basis in the Hilbert space of the considered system, and the eigenstates of the total Hamiltonian can thus be written as

$$|\Psi\rangle = \sum_n \alpha_n |\{n\}\rangle. \quad (\text{I.12})$$

The solution will be exact if the set of MOs forms a complete basis of single-particle states and the many-particle wave function includes all Slater determinants (of which there are $(M N)$). This is then called the FCI (*full configuration interaction*) wave function.

As already mentioned, in the formalism of second quantization, the antisymmetry property is transferred from Slater's determinant to the algebraic properties of certain operators. These operators excite ("create") and de-excite (annihilate) electrons from individual MOs. The action of the creation operator a_p^\dagger can be perceived as adding a particle in the state $|\phi_p\rangle$ to the right of the already existing states

$$a_p^\dagger |n_{M-1}, \dots, 0_p, \dots, n_0; A\rangle = \sqrt{N!} A (|\phi_{M-1}\rangle^{\otimes n_{M-1}} \otimes \dots \otimes |\phi_0\rangle^{\otimes n_0}) |\phi_p\rangle = (-1)^{\sum_{i=0}^{p-1} n_i} |n_{M-1}, \dots, 1_p, \dots, n_0; A\rangle, \quad (\text{I.13})$$

where the sum counts over occupied states and we define

$$a_p^\dagger |n_{M-1}, \dots, 1_p, \dots, n_0; A\rangle = 0. \quad (\text{I.14})$$

The annihilation operator is given by the action

$$a_p |n_{M-1}, \dots, 1_p, \dots, n_0; A\rangle = (-1)^{\sum_{i=0}^{p-1} n_i} |n_{M-1}, \dots, 0_p, \dots, n_0; A\rangle \quad (\text{I.15})$$

and we define

$$a_p |n_{M-1}, \dots, 0_p, \dots, n_0; A\rangle = 0. \quad (\text{I.16})$$

It can be shown that the requirement of antisymmetrization of fermion states implies a condition for fermion creation and annihilation operators in the form of the satisfying the anticommutation relations

$$\{a_p, a_q\} = \{a_p^\dagger, a_q^\dagger\} = 0 \quad \{a_p, a_q^\dagger\} = \delta_{pq} \quad (\text{I.17})$$

Observables can be rewritten using the creation and annihilation operators. Let O be a single-particle observable. In the basis of one-particle states $\{\phi_j\}$ can be written as

$$O = \sum_{p,q} \langle \phi_p | O | \phi_q \rangle | \phi_p \rangle \langle \phi_q |. \quad (\text{I.18})$$

For N particles we can write

$$O^{(1)} = \sum_{k=0}^N I_{M-1} \otimes \dots \otimes I_{k+1} \otimes O \otimes I_{k-1} \otimes \dots \otimes I_0. \quad (\text{I.19})$$

Thus for particles in states $|\phi_{k_0}\rangle \dots |\phi_{k_{N-1}}\rangle$

$$O^{(1)} |\phi_{k_{N-1}} \dots \phi_{q_i} \dots \phi_{k_0}; A\rangle = \sum_{k=0}^N \sum_{p_i} \langle \phi_{p_i} | O | \phi_{q_i} \rangle |\phi_{k_{N-1}} \dots \phi_{p_i} \dots \phi_{k_0}; A\rangle \quad (\text{I.20})$$

and using

$$|\phi_{k_{N-1}} \dots \phi_{k_0}; A\rangle = a_{k_{N-1}}^\dagger \dots a_{k_0}^\dagger |0\rangle \quad (\text{I.21})$$

and commutation relations, it is then easy to show that the one-particle operator can be written using the creation and annihilation operators as

$$O^{(1)} = \sum_{p,q} \underbrace{\langle \phi_p | O | \phi_q \rangle}_{h_{pq}} a_p^\dagger a_q = \sum_{p,q} h_{pq} a_p^\dagger a_q, \quad (\text{I.22})$$

where

$$h_{pq} = \int d\mathbf{x} \phi_p^*(\mathbf{x}) O \phi_q(\mathbf{x}). \quad (\text{I.23})$$

In a similar way, it can be deduced that the two-particle operator can be written using the creation and annihilation operators as

$$O^{(2)} = \frac{1}{2} \sum_{p,q,r,s} \underbrace{\langle \phi_p \phi_q | O | \phi_r \phi_s \rangle}_{h_{pqrs}} a_p^\dagger a_q^\dagger a_r a_s = \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \quad (\text{I.24})$$

where

$$h_{pqrs} = \int d\mathbf{x}_1 d\mathbf{x}_2 \phi_p^*(\mathbf{x}_1) \phi_q^*(\mathbf{x}_2) O(\phi_r(\mathbf{x}_2) \phi_s(\mathbf{x}_1)). \quad (\text{I.25})$$

The electron Hamiltonian I.3 is thus obtained in the formulation of the second quantization as

$$H_e = \sum_{p,q} h_{pq} a_p^\dagger a_q + \frac{1}{2} \sum_{p,q,r,s} h_{pqrs} a_p^\dagger a_q^\dagger a_r a_s, \quad (\text{I.26})$$

where

$$h_{pq} = \int d\mathbf{x} \phi_p^*(\mathbf{x}) \left(-\frac{\hbar^2}{2m_e} \nabla^2 - \sum_I \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{r} - \mathbf{R}_I|} \right) \phi_q(\mathbf{x}) \quad (\text{I.27})$$

and

$$h_{pqrs} = \int d\mathbf{x}_1 d\mathbf{x}_2 \frac{\phi_p^*(\mathbf{x}_1) \phi_q^*(\mathbf{x}_2) \phi_r(\mathbf{x}_2) \phi_s(\mathbf{x}_1)}{|\mathbf{r}_1 - \mathbf{r}_2|}. \quad (\text{I.28})$$

I.3 Born-Oppenheimer approximation

Born-Oppenheimer approximation (BOA) [11] treats nuclei as point charges and separates their motion from the motion of electrons [36]. This is made possible by the fact that the mass of nuclei is three orders of magnitude greater than the mass of electrons. Therefore, BOA transforms the problem I.1 into the motion of electrons in the stationary potential of nuclei. The many-particle wave function appearing in I.1 is a function of the coordinates of electrons and nuclei $|\Psi(\mathbf{r}, \mathbf{R})\rangle$. By fixing the positions of the nuclei, we obtain an electronic Hamiltonian I.3 parameterized by them, where this parameterization stems from the term

$$V_{en}(\mathbf{R}) = - \sum_{i,I} \frac{e^2}{4\pi\epsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|}. \quad (\text{I.29})$$

Solving the electron Schrödinger equation

$$H_e(\mathbf{R}) |\psi_k^{\mathbf{R}}(\mathbf{r})\rangle = E_{e,k}^{\mathbf{R}} |\psi_k^{\mathbf{R}}(\mathbf{r})\rangle \quad (\text{I.30})$$

will give us a complete basis of eigenstates for electrons with fixed nuclear positions $|\psi_k^{\mathbf{R}}(\mathbf{r})\rangle$. The solution of the Schrödinger equation for the entire system can then be written in the form

$$|\Psi_n(\mathbf{r}, \mathbf{R})\rangle = \sum_k |\varphi_{nk}(\mathbf{R})\rangle |\psi_k^{\mathbf{R}}(\mathbf{r})\rangle, \quad (\text{I.31})$$

where $\varphi_{nk}(\mathbf{R})$ are nuclear functions that here serve as expansion coefficients.

Now, an approximation will be made, due to which BOA is sometimes also called the adiabatic approximation. It will be considered that nuclei move so slowly that electrons react to this movement "adiabatically." This means that when the coordinates of the nuclei change, the electronic wave function changes only its parameterization and not its index indicating the solution of equation I.30. In other words, a change $\mathbf{R} \rightarrow \mathbf{R}'$ leads only to a change from $\psi_k^{\mathbf{R}}(\mathbf{r}) \rightarrow \psi_k^{\mathbf{R}'}(\mathbf{r})$, and the index k remains preserved.

This is important because, during the time evolution, if we allow the nuclei to move, this movement will not affect the k -state of the electronic wave function. Therefore, the stationary state of the entire system can be related to a single k -state and written as

$$|\Psi_n(\mathbf{r}, \mathbf{R})\rangle = |\varphi_{nk}(\mathbf{R})\rangle |\psi_k^{\mathbf{R}}(\mathbf{r})\rangle. \quad (\text{I.32})$$

By explicitly substituting I.31 into I.1 we get

$$\underbrace{(T_n + V_{nn} + H_e)}_{H_{mol}} \sum_{k'} |\varphi_{nk'}(\mathbf{R})\rangle |\psi_{k'}^{\mathbf{R}}(\mathbf{r})\rangle = E_n \sum_{k'} |\varphi_{nk'}(\mathbf{R})\rangle |\psi_{k'}^{\mathbf{R}}(\mathbf{r})\rangle. \quad (\text{I.33})$$

Multiplying by $\langle \psi_k^{\mathbf{R}}(\mathbf{r}) |$ from the left goes to

$$\begin{aligned} \sum_{k'} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | T_n + \underbrace{V_{nn}}_{\text{can be factored out}} | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle |\varphi_{nk'}(\mathbf{R})\rangle &+ \sum_{k'} \underbrace{\langle \psi_k^{\mathbf{R}}(\mathbf{r}) | H_e | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle}_{E_{e,k'} \delta_{kk'}} |\varphi_{nk'}(\mathbf{R})\rangle \\ &= E_n \sum_{k'} \underbrace{\langle \psi_k^{\mathbf{R}}(\mathbf{r}) | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle}_{\delta_{kk'}} |\varphi_{nk'}(\mathbf{R})\rangle. \end{aligned} \quad (\text{I.34})$$

The only term that needs a closer look is the kinetic energy term of the nuclei. Its action is different for different k -states and represents precisely the response of electrons to the movement of nuclei.

$$\begin{aligned} \sum_{k'} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | T_n | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle |\varphi_{nk'}(\mathbf{R})\rangle &= - \sum_{k'} \sum_I \frac{\hbar^2}{2M_I} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | \nabla_I^2 | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle |\varphi_{nk'}(\mathbf{R})\rangle = \\ &= \sum_{k'} \underbrace{\sum_I -\frac{\hbar^2}{2M_I} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | \nabla_I^2 | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle}_{B_{kk'}} |\varphi_{nk'}(\mathbf{R})\rangle + \sum_{k'} \underbrace{\sum_I -\frac{\hbar^2}{M_I} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | \nabla_I | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle}_{A_{kk'}} \nabla_I |\varphi_{nk'}(\mathbf{R})\rangle - \\ &\quad - \sum_{k'} \underbrace{\sum_I \frac{\hbar^2}{2M_I} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | \psi_{k'}^{\mathbf{R}}(\mathbf{r}) \rangle}_{\delta_{kk'}} \nabla_I^2 |\varphi_{nk'}(\mathbf{R})\rangle = T_n |\varphi_{nk}(\mathbf{R})\rangle + \sum_{k'} (A_{kk'} + B_{kk'}) |\varphi_{nk'}(\mathbf{R})\rangle \end{aligned} \quad (\text{I.35})$$

By substituting I.35 into I.34 and rearranging, we get

$$[T_n + V_{nn} - E_n + E_{e,k}] |\varphi_{nk}(\mathbf{R})\rangle = - \sum_{k'} (A_{kk'} + B_{kk'}) |\varphi_{nk'}(\mathbf{R})\rangle. \quad (\text{I.36})$$

It is easy to show that due to normalization the diagonal term $A_{kk} = 0$:

$$\begin{aligned} A_{kk} &= \sum_I -\frac{\hbar^2}{2M_I} \langle \psi_k^{\mathbf{R}}(\mathbf{r}) | \nabla_I | \psi_k^{\mathbf{R}}(\mathbf{r}) \rangle \nabla_I = \sum_I -\frac{\hbar^2}{2M_I} \int (\psi_k^{\mathbf{R}}(\mathbf{r}))^* \nabla_I \psi_k^{\mathbf{R}}(\mathbf{r}) d\mathbf{r} \nabla_I = \\ &= \sum_I -\frac{\hbar^2}{2M_I} \frac{1}{2} \nabla_I \underbrace{\int \psi_k^{\mathbf{R}}(\mathbf{r}) \psi_k^{\mathbf{R}}(\mathbf{r}) d\mathbf{r}}_1 \nabla_I = 0, \end{aligned} \quad (\text{I.37})$$

where in the third equality, the fact that for stationary states it is possible to choose the function $\psi_k^{\mathbf{R}}(\mathbf{r})$ real was used. If we also separate the diagonal term B_{ii} (which does not change the k -state due to the movement of the nucleus), I.36 can be written as

$$[T_n + V_{nn} - E_n + E_{e,k} + B_{kk}] |\varphi_{nk}(\mathbf{R})\rangle = - \sum_{k', k' \neq k} (A_{kk'} + B_{kk'}) |\varphi_{nk'}(\mathbf{R})\rangle. \quad (\text{I.38})$$

To perform BOA then means to set the sum over the off-diagonal terms equal to 0. It is precisely these terms that mix the individual k -states during the movement of the nuclei. The equation after performing BOA takes the form

$$[T_n + V_{nn} - E_n + E_{e,k} + B_{kk}] |\varphi_{nk}(\mathbf{R})\rangle = 0. \quad (\text{I.39})$$

For each electronic k -state, we thus obtain an equation only for the nuclei.

BOA ceases to be valid if a given electronic level is degenerate or if the difference between the ground energy level (which we are primarily interested in) and the excited levels is so small that it is of the order of the energy of nuclear motion. Usually, however, the difference between the levels of electronic k -states is much larger, and thus BOA approximates very well. Next, we will focus on calculating the eigenvalues and eigenstates (primarily the ground state) of equation I.30. We will thus be addressing the so-called problem of electronic structure.

I.4 Hartree-Fock method

The Hartree-Fock method is based on Hartree's method [25] and provides a way to find the dominant Slater determinant. Thus, it makes an approximation by assuming that the complete N -particle wave function can be expressed by just one Slater determinant. It is based on the variational principle, where the resulting wave function with the lowest expected energy is achieved through iterative optimization of MOs. The derivation will be performed in the language of first quantization and subsequently indicated how it would be equivalently proceeded in the formulation of second quantization.

I.4.1 Variational principle

Let H be a Hamiltonian with eigenvalues E_μ and eigenstates $|\mu\rangle$ and let $|\psi_0\rangle$ and E_0 be the ground state and ground energy, respectively. We assume that the eigenstates of $|\mu\rangle$ form a complete orthonormal basis, and we can therefore write any function as

$$|\psi\rangle = \sum_\mu a_\mu |\mu\rangle. \quad (\text{I.40})$$

For the expected value of the energy then

$$\tilde{E}(|\psi\rangle) = \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle} = \frac{\sum_\mu E_\mu |a_\mu|^2}{\sum_\mu |a_\mu|^2} = E_0 + \frac{\sum_\mu (E_\mu - E_0) |a_\mu|^2}{\sum_\mu |a_\mu|^2} \quad (\text{I.41})$$

and since $E_0 \leq E_\mu$, then $\tilde{E}(|\psi\rangle) \geq E_0$ must hold. This gives us the basic equation of the variation principle

$$E_0 \leq \frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}. \quad (\text{I.42})$$

Next, we parametrize $|\psi\rangle$ with a set of parameters \mathbf{p} and require

$$\frac{\partial \tilde{E}(|\psi(\mathbf{p})\rangle)}{\partial \mathbf{p}} = 0, \quad (\text{I.43})$$

i.e., we require stationarity of \tilde{E} with respect to variations of $|\psi\rangle$. However, for this the normalization condition $\langle\psi|\psi\rangle = 1$ must hold, which leads to the theory of constrained extrema. It is known from that

that if we are looking for the minimum of some function $f(\mathbf{r})$ under the condition $g(\mathbf{r}) = 0$, then the condition for the minimum is given as

$$\nabla_{\mathbf{r},\lambda} \left(\underbrace{f(\mathbf{r}) - \lambda g(\mathbf{r})}_L \right) = \nabla_{\mathbf{r},\lambda} L(\mathbf{r}, \lambda) = 0, \quad (\text{I.44})$$

where L is called the Lagrange function and λ is the Lagrange multiplier. In our case

$$L = \langle \psi | H | \psi \rangle - \lambda \langle \psi | \psi \rangle, \quad (\text{I.45})$$

which leads to the condition

$$\delta L = \delta [\langle \psi | H | \psi \rangle - \lambda \langle \psi | \psi \rangle] = 0. \quad (\text{I.46})$$

I.4.2 Hartree-Fock equations

The Hartree-Fock method is based on the Slater determinant I.5, which, however, is convenient to rewrite in the form

$$|\psi(\mathbf{x}_0, \mathbf{x}_1, \dots, \mathbf{x}_{N-1})\rangle = \frac{1}{\sqrt{N!}} \sum_{\pi} (-1)^{\pi} P_{\pi} |\phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1})\rangle, \quad (\text{I.47})$$

where $(-1)^{\pi}$ denotes the sign of the permutation. The variational principle then enforces

$$\delta(\langle \psi | H | \psi \rangle - \sum_{i,j} \lambda_{ij} \langle \phi_i | \phi_j \rangle) = 0. \quad (\text{I.48})$$

Before analyzing the individual terms, it is convenient to rewrite the Hamiltonian I.3 as

$$H_e \equiv H = \underbrace{T_e + V_{en}}_{\text{single-particle}} + \underbrace{V_{ee}}_{\text{two-particle}} = \sum_i (T_i + \underbrace{\sum_I V_{iI}}_{H_i}) + \frac{1}{2} \sum_{i,j, i \neq j} V_{ij} = \sum_i H_i + \frac{1}{2} \sum_{i,j, i \neq j} V_{ij}. \quad (\text{I.49})$$

The following is a term by term approach, only one term will be explicitly resolved here.

$$\langle \psi | \sum_i H_i | \psi \rangle = \frac{1}{N!} \sum_{\pi, \pi'} (-1)^{\pi+\pi'} \left(P_{\pi} \langle \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) | \right) \sum_i H_i \left(P_{\pi'} | \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) \rangle \right) \quad (\text{I.50})$$

Factoring out P_{π} and denoting $P_{\sigma} = P_{\pi'} P_{\pi}^{-1}$ we get

$$\langle \psi | \sum_i H_i | \psi \rangle = \frac{1}{N!} \sum_{\pi, \pi'} (-1)^{\sigma} P_{\pi} \left(\langle \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) | \sum_i H_i P_{\sigma} | \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) \rangle \right) \quad (\text{I.51})$$

and since the permutation P_{π} produces $N!$ identical integrals, we can write

$$\begin{aligned} \langle \psi | \sum_i H_i | \psi \rangle &= \sum_{\sigma} (-1)^{\sigma} \langle \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) | \sum_i H_i P_{\sigma} | \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) \rangle = \\ &= \sum_{\sigma} (-1)^{\sigma} \langle \phi_0(\mathbf{x}_0) \dots \phi_{N-1}(\mathbf{x}_{N-1}) | \sum_i H_i | \phi_{\sigma(0)}(\mathbf{x}_0) \dots \phi_{\sigma(N-1)}(\mathbf{x}_{N-1}) \rangle = \\ &= \sum_{\sigma} (-1)^{\sigma} \sum_i \langle \phi_i(\mathbf{x}_i) | H_i | \phi_{\sigma(i)}(\mathbf{x}_i) \rangle \delta_{1,\sigma(1)} \dots \delta_{i-1,\sigma(i-1)} \delta_{i+1,\sigma(i+1)} \dots \delta_{N-1,\sigma(N-1)} = \\ &= \sum_i \langle \phi_i(\mathbf{x}_0) | H_0 | \phi_i(\mathbf{x}_0) \rangle, \end{aligned} \quad (\text{I.52})$$

where the orthogonality of one-particle functions and the fact that H_i acts only on the space \mathbf{x}_i were used in the third equality, and until the last equality only an identical permutation remained and the integration variable was relabeled, since all spaces are the same.

An analogous procedure [45] for the two-particle part of the Hamiltonian yields

$$\langle \psi | \frac{1}{2} \sum_{i,j,i \neq j} V_{ij} | \psi \rangle = \frac{1}{2} \sum_{i,j,i \neq j} \left(\underbrace{\left(\langle \phi_i(\mathbf{x}_0) \phi_j(\mathbf{x}_1) | V_{12} | \phi_i(\mathbf{x}_0) \phi_j(\mathbf{x}_1) \rangle - \langle \phi_i(\mathbf{x}_0) \phi_j(\mathbf{x}_1) | V_{12} | \phi_j(\mathbf{x}_0) \phi_i(\mathbf{x}_1) \rangle \right)}_{\text{Hartree term}} \right. \\ \left. - \underbrace{\left(\langle \phi_i(\mathbf{x}_0) \phi_j(\mathbf{x}_1) | V_{12} | \phi_j(\mathbf{x}_0) \phi_i(\mathbf{x}_1) \rangle - \langle \phi_j(\mathbf{x}_0) \phi_i(\mathbf{x}_1) | V_{12} | \phi_i(\mathbf{x}_0) \phi_j(\mathbf{x}_1) \rangle \right)}_{\text{exchange term}} \right). \quad (\text{I.53})$$

Finally, the variation principle I.48 can be written in the form

$$\delta_k (\langle \psi | H | \psi \rangle - \sum_i \lambda_i \langle \phi_i | \phi_i \rangle) = 0, \quad (\text{I.54})$$

where δ_k is a variation of the function ϕ_k and thus $\delta_k \phi_i = 0$ for $i \neq k$. By rewriting

$$\delta_k \left(\langle \psi | \sum_i H_i | \psi \rangle \right) = \langle \delta_k \psi | \sum_i H_i | \psi \rangle + \langle \psi | \sum_i H_i | \delta_k \psi \rangle = \langle \delta_k \psi | \sum_i H_i | \psi \rangle + \left(\langle \delta_k \psi | \sum_i H_i | \psi \rangle \right)^* = \\ = 2 \operatorname{Re} [\langle \delta_k \phi_k(\mathbf{x}_0) | H_0 | \phi_k(\mathbf{x}_0) \rangle] \quad (\text{I.55})$$

and similarly

$$\delta_k \langle \psi | \frac{1}{2} \sum_{i,j,i \neq j} V_{ij} | \psi \rangle = \operatorname{Re} [\langle \delta_k \phi_k(\mathbf{x}_0) | \left(\sum_{j,j \neq k} \langle \phi_j(\mathbf{x}_1) | V_{12} | \phi_j(\mathbf{x}_1) \rangle | \phi_k(\mathbf{x}_0) \rangle - \right. \\ \left. - \sum_{j,j \neq k} \langle \phi_j(\mathbf{x}_1) | V_{12} | \phi_k(\mathbf{x}_1) \rangle | \phi_j(\mathbf{x}_0) \rangle \right)] \quad (\text{I.56})$$

and for the normalization condition

$$\delta (\langle \phi_k | \phi_k \rangle) = 2 \operatorname{Re} [\langle \delta_k \phi_k | \phi_k \rangle]. \quad (\text{I.57})$$

Together, the variational principle gives

$$\operatorname{Re} [\langle \delta_k \phi_k(\mathbf{x}_0) | \left(H_0 | \phi_k(\mathbf{x}_0) \rangle + \sum_{j,j \neq k} \langle \phi_j(\mathbf{x}_1) | V_{12} | \phi_j(\mathbf{x}_1) \rangle | \phi_k(\mathbf{x}_0) \rangle - \right. \\ \left. - \sum_{j,j \neq k} \langle \phi_j(\mathbf{x}_1) | V_{12} | \phi_k(\mathbf{x}_1) \rangle | \phi_j(\mathbf{x}_0) \rangle - \lambda_k | \phi_k \rangle \right)] = 0. \quad (\text{I.58})$$

From this, simply renaming the index k to i and renaming the Lagrange multiplier λ_i to ε_i results in a set of coupled equations called the Hartree-Fock equations

$$H_0 | \phi_i(\mathbf{x}_0) \rangle + \sum_j \langle \phi_j(\mathbf{x}_1) | V_{12} | \phi_j(\mathbf{x}_1) \rangle | \phi_i(\mathbf{x}_0) \rangle - \sum_j \langle \phi_j(\mathbf{x}_1) | V_{12} | \phi_i(\mathbf{x}_1) \rangle | \phi_j(\mathbf{x}_0) \rangle = \varepsilon_i | \phi_i(\mathbf{x}_0) \rangle, \quad (\text{I.59})$$

where it was utilized that the terms $j = i$ cancel each other out from both sums, thus allowing summation over all j . It is evident that the Lagrange multipliers take on the meaning of energy, and from Koopmans' theorem, it follows that this is the energy required to remove an electron, assuming that the state of the other electrons does not change. This interpretation, therefore, is an approximation that may not always be valid.

As already mentioned, a derivation in second quantization will also be indicated. This itself does not introduce any new concept, and therefore the details of the computation will be omitted. The starting

point is the Hamiltonian I.26, hereafter simply denoted as H . We assume that the occupied states of the normalized ground state minimizing the expected energy are labeled $0, \dots, N - 1$. For the state $|\psi\rangle$ providing the minimum energy, the from the variational principle it must hold that

$$\langle \delta\psi | H | \psi \rangle = 0. \quad (\text{I.60})$$

$|\delta\psi\rangle$ can be thought of, for example, as

$$|\delta\psi\rangle = (\delta s) a_{k'}^\dagger a_{i'} |\psi\rangle, \quad (\text{I.61})$$

where δs is a small number, $i' \leq N - 1$ ($a_{i'}$ annihilates an electron from the ground state) and $k' > N - 1$ ($a_{k'}^\dagger$ creates an electron to a state above the ground state). It is not difficult to show that in this way the state $|\psi\rangle + |\delta\psi\rangle$ is also normalized to the first order of variation. The variational principle thus gives

$$0 = \langle \delta\psi | H | \psi \rangle = (\delta s)^* \langle \psi | H a_{i'}^\dagger a_{k'} | \psi \rangle \quad (\text{I.62})$$

and substituting into I.26

$$0 = \sum_{i,j} h_{ij} \langle \psi | a_{i'}^\dagger a_{k'} a_i^\dagger a_j | \psi \rangle + \frac{1}{2} \sum_{i,j,k,l} h_{ijkl} \langle \psi | a_{i'}^\dagger a_{k'} a_i^\dagger a_j^\dagger a_k a_l | \psi \rangle, \quad (\text{I.63})$$

where i, j, k, l sums over occupied and unoccupied MOs. Using anticommutation relations and the action of operators, the first term I.63 can be written as

$$\langle \psi | a_{i'}^\dagger a_{k'} a_i^\dagger a_j | \psi \rangle = \delta_{i',j} \delta_{k',i} \quad (\text{I.64})$$

and the second as

$$\langle \psi | a_{i'}^\dagger a_{k'} a_i^\dagger a_j^\dagger a_k a_l | \psi \rangle = -\delta_{k',j} \delta_{i',l} \delta_{i,k \leq N-1} + \delta_{k',j} \delta_{i,l \leq N-1} \delta_{i',k} + \delta_{k',i} \delta_{i',l} \delta_{j,k \leq N-1} - \delta_{k',j} \delta_{j,l \leq N-1} \delta_{i',k}. \quad (\text{I.65})$$

By the action of Kronecker δ and using the symmetry $h_{ijkl} = h_{jilk}$ it is possible to write

$$0 = h_{k'i'} + \sum_{j=0}^{N-1} (h_{k'jj'i} - h_{k'ji'j}) \quad (\text{I.66})$$

This already leads to the introduction of the so-called Fock operator f as

$$f = \sum_{i,j} (h_{ij} + V_{ij}) a_i^\dagger a_j, \quad (\text{I.67})$$

where

$$V_{ij} = \sum_{k=0}^{N-1} (h_{ikkj} - h_{ikjk}). \quad (\text{I.68})$$

Apparently, the h_{ikkj} term describes the Hartree (Coulomb) interaction, while the h_{ikjk} term accounts for the exchange energy. The Fock operator is Hermitian and thus diagonalizable, i.e. can be written as

$$f = \sum_i \varepsilon_i a_i^\dagger a_i. \quad (\text{I.69})$$

The Hartree-Fock equations follow from this in the form

$$h_{ij} + V_{ij} = \varepsilon_j \delta_{ij}. \quad (\text{I.70})$$

I.4.3 Numerical calculation, basis sets

For better clarity, we will continue in the formulation of first quantization, where it can be noticed that multiplying the H-F equation I.59 from the left by $\langle \phi_j(\mathbf{x}_0) |$, we precisely get equations I.70 that emerged in second quantization. Thus, on the left side of I.59 is nothing other than the Fock operator. For simplicity, let's assume that N is even. It's then straightforward to show that the Fock operator is spin-independent, which effectively means we are looking for $N/2$ spatial states, each accommodating 2 electrons with opposite spins. If the single-particle function has the form $|\phi_i(\mathbf{x}_0)\rangle = |\phi_i(\mathbf{r}_0)\rangle |\xi_i(\sigma)\rangle$, $\sigma = \pm 1/2$, then we solve $N/2$ coupled equations of the form

$$H_0 |\phi_i(\mathbf{r}_0)\rangle + \sum_j \langle \phi_j(\mathbf{r}_1) | V_{12} |\phi_j(\mathbf{r}_1)\rangle |\phi_i(\mathbf{r}_0)\rangle - \sum_j \langle \phi_j(\mathbf{r}_1) | V_{12} |\phi_i(\mathbf{r}_1)\rangle |\phi_j(\mathbf{r}_0)\rangle = \varepsilon_i |\phi_i(\mathbf{r}_0)\rangle. \quad (\text{I.71})$$

This set consists of a series of nonlinear integral equations that are solved iteratively. In the k -th step, a set of MO $|\phi_i(\mathbf{r}_0)\rangle^k$ is obtained. With these, the Coulomb and exchange terms are calculated and then used in the $k+1$ -th iteration to obtain a new set $|\phi_i(\mathbf{r}_0)\rangle^{k+1}$. This process repeats until the MO converge. Hence, the Hartree-Fock method is also referred to as the self-consistent field (SCF) method. Unfortunately, even for numerical calculations, solving this set of nonlinear integral equations is too complex. The solution, introduced by C.C.J. Roothaan [53], is to transform these equations (under certain approximations) into algebraic equations. The fundamental idea is that MO can approximately be expressed as a linear combination of some basis set (where, in computational chemistry terminology, a basis set indeed refers to a collection of functions used in the linear combination to model MO).

$$|\phi_i\rangle \approx \sum_{k=0}^{L-1} c_{ki} |\chi_k\rangle, \quad (\text{I.72})$$

where χ_k is a basis set of square-integrable functions, and for implementation purposes, we select some L -element subset from it (it's worth noting that to subsequently obtain square matrices, it's necessary to have $L \geq N/2$). These functions do not necessarily need to be orthonormal. The choices for basis functions will be discussed shortly, but first, the algebraic equations will be derived in general. The equation I.71 formally has the form

$$f(\mathbf{r}_0) |\phi_i(\mathbf{r}_0)\rangle = \varepsilon_i |\phi_i(\mathbf{r}_0)\rangle, \quad (\text{I.73})$$

where f is the Fock operator. Substituting I.72 into I.73 and multiplying $\langle \chi_j(\mathbf{r}_0) |$ from the left, we get

$$\sum_k c_{ki} \underbrace{\langle \chi_j(\mathbf{r}_0) | f(\mathbf{r}_0) | \chi_k(\mathbf{r}_0)\rangle}_{F_{jk}} = \varepsilon_i \sum_k c_{ki} \underbrace{\langle \chi_j(\mathbf{r}_0) | \chi_k(\mathbf{r}_0)\rangle}_{S_{jk}}. \quad (\text{I.74})$$

Toto lze maticově zapsat jako

$$\mathbb{F}\mathbb{C} = \mathbb{S}\mathbb{C}\mathbb{E}, \quad (\text{I.75})$$

which is called the Hartree-Fock-Roothaan equation for the unknown coefficient matrix \mathbb{C} . Here, $\mathbb{E} = \text{diag}(\varepsilon_i)$ contains single-particle energies on its diagonal, \mathbb{F} denotes the Fock matrix, and \mathbb{S} represents the overlap matrix. The coefficients in the n -th column of the matrix \mathbb{C} represent the expansion coefficients of MO ϕ_n in the basis $\{\chi_k\}_{k=0}^{L-1}$. Thus, it represents a generalized eigenvalue problem (apparent when pre-multiplied by \mathbb{S}^{-1} from the left, although this is generally not done in the actual solution process), where the columns of the matrix \mathbb{C} form the eigenvectors and ε_i the eigenvalues. As mentioned with the Hartree-Fock equation, this equation is solved iteratively, where in the k -th step, the elements of the matrix \mathbb{C}_k are computed. These are then used to calculate the elements of the matrix \mathbb{F}_k , which is

subsequently substituted back into I.75 to obtain $\mathbb{C}k + 1$. This process repeats until the elements of this matrix converge.

In the following paragraphs, the choices of basis sets will be briefly discussed. A good basis set should fulfill two properties. The first is that it should well capture the physics of the problem, effectively meaning that as few basis functions as possible are needed for an adequately good approximation of the actual MO. The second is that the basis functions should have a suitable mathematical form to facilitate the computation of molecular integrals. Different basis sets offer various degrees of compromise between these properties.

Broadly, molecular systems can be divided into periodic systems and isolated molecules. Molecular crystals or large biomolecules can be considered an intermediate stage between these extremes. For this work, isolated molecules are more crucial, so only a few remarks will be devoted to periodic systems. In periodic systems, especially in metals, valence electrons behave almost like free electrons, hence, plane waves are a straightforward choice for the basis set.

$$\chi_k(\mathbf{r}) = N_{PW} \exp\{i\mathbf{G}_k \cdot \mathbf{r}\}, \quad (\text{I.76})$$

where \mathbf{G}_k is the reciprocal lattice vector and N_{PW} is the normalization constant. However, a large number of these waves are needed to describe the inner (core) electrons, and therefore pseudopotential methods are used. This also shows that plane waves are not ideal for describing isolated molecules.

In isolated molecules, the electron density is maximal around the nuclei and decreases exponentially with distance from the nucleus. This automatically leads to the idea that the MO could be expressed in a basis of atomic orbitals (AO) arising from the solution of the Schrödinger equation for atomic hydrogen. Such solutions have the form

$$\chi_k(\mathbf{r}) = R_{k,nl}(r) Y_{lm}(\theta, \varphi), \quad (\text{I.77})$$

where n is the principal quantum number, l denotes the orbital angular momentum quantum number, and m is the magnetic quantum number. $Y_{lm}(\theta, \varphi)$ is the spherical function and $R_{nl}(r)$ is the radial part of the function, which is the product of a Laguer polynomial and a term decaying exponentially with r . However, the solutions obtained in this way dissipate too quickly for a many-electron problem. The so-called Slater-type orbitals (STO), which are of the form

$$\chi_k(\mathbf{r}) = N_{STO} \underbrace{(\zeta_k r)^{n-1} \exp\{-\zeta_k r\}}_{R_n^{STO}(r)} Y_{lm}(\theta, \varphi), \quad (\text{I.78})$$

where N_{STO} is the normalization constant and ζ is a fitting parameter that controls how quickly the probability density of the given orbital disperses with distance from the nucleus. Unlike real orbitals, these functions do not oscillate sharply around the nucleus, hence a linear combination of them is needed for an accurate description. However, they exhibit the same exponential decay with r . If we consider for each MO only one basis function with some parameter ζ , it's called a single-zeta representation. Conversely, if each MO is represented by n STOs with different ζ , it's called an n -zeta representation. The advantage of STOs is the correct capture of electron behavior close to and far from the nucleus. Unfortunately, the computation of molecular integrals with them is not straightforward and must be solved numerically, which limits STOs to small molecules.

This motivates the introduction of Gaussian-type orbitals (GTOs), which take the form

$$\chi_k(\mathbf{r}) = N_{GTO} \underbrace{(\sqrt{\alpha_k} r)^l \exp\{-\alpha_k r^2\}}_{R_{nl}^{GTO}(r)} Y_{lm}(\theta, \varphi). \quad (\text{I.79})$$

Here, α_k is again a fitting parameter. GTOs have excellent analytical properties that allow for efficient evaluation of molecular integrals. Unlike STOs, however, due to the exponential decay with r^2 , they do

not precisely capture the dispersion of electron density. Somewhere between accuracy and computational efficiency lies today's most commonly used method, where one STO is expressed (approximately) by a linear combination of several GTOs. These STOs are then used as the basis set. There are many combinations, for example, where valence electrons are assigned two STOs (double-zeta), while core electrons are assigned only one, or they are assigned so-called diffuse functions (GTOs with a very small parameter). The number of basis functions then affects the length and memory requirements of the calculation in classical computation. In quantum computation, this determines the number of qubits and gates used.

I.5 Post-Hartree-Fock methods, typical calculation

The Hartree-Fock (H-F) method provides a Slater determinant composed of Molecular Orbitals (MOs) corresponding to the ground state. While it accurately reflects the fermionic nature of electrons, it neglects the energy arising from electron correlations. These can be broadly divided into static (contributions from bond dissociations, excitations, etc.) and dynamic (contributions from Coulomb repulsion) correlations. Static correlations can be accounted for by including additional determinants containing MOs corresponding to excitations above the H-F state. However, it must be noted that the H-F method does not provide any good representation for excitation orbitals, optimizing only the MOs occupied in that single Slater determinant. Dynamic correlation functions then serve to eliminate the effects of the mean-field approximation used in the H-F method. Today, the state provided by the H-F method is taken as the reference state for so-called post-H-F methods. Some of these will be briefly mentioned here.

The Configuration Interaction (CI) method provides a correlated wave function by including excitations above some reference state - usually the H-F state. If we were to consider all possible excitations and their combinations (i.e., all possible Slater determinants), we would obtain full configuration interaction (FCI).

$$|\psi_{FCI}\rangle = \left(I + \sum_{i,\alpha} C_{ia} a_i^\dagger a_\alpha + \sum_{i>j,\alpha>\beta} C_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta + \dots \right) |\psi_{HF}\rangle, \quad (\text{I.80})$$

where C are parameters optimized using the variation principle. To make this method computationally tractable, usually only a few excitations above the H-F state are considered.

A similar approach is chosen by the coupled cluster (CC) method, where, however, a product scheme is chosen for parameterization

$$|\psi_{CC}\rangle = \prod_{i,\alpha} \left(I + C_{ia} a_i^\dagger a_\alpha \right) \cdot \prod_{i>j,\alpha>\beta} \left(I + C_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta \right) \cdot \dots |\psi_{HF}\rangle. \quad (\text{I.81})$$

This can be rewritten as

$$|\psi_{CC}\rangle = e^T |\psi_{HF}\rangle, \quad (\text{I.82})$$

where $T = \sum_i T_i$,

$$T_1 = \sum_{i \in ex, \alpha \in occ} t_{ia} a_i^\dagger a_\alpha \quad T_2 = \sum_{i,j \in ex, \alpha, \beta \in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta \quad T_3 = \dots \quad (\text{I.83})$$

Here, occ denotes the orbitals occupied in the H-F state, and ex represents the unoccupied ones. Again, if we consider all T_i , we obtain FCI. This method provides (among other advantages) faster convergence than CI; however, the resulting function does not adhere to the variational principle. Moreover, this method takes only a single Slater determinant as input, making it unsuitable for systems with strong static correlations described by a superposition of several determinants. The modified method, unitary coupled

cluster (UCC), then addresses both these issues but is effectively implementable only on a quantum computer.

Finally, let's mention the multiconfigurational self-consistent field (MCSCF) method, which considers a wave function composed of several Slater determinants and variationally optimizes both the MOs and the amplitudes of individual determinants. It is therefore suitable for calculations on systems with strong static correlations, where several determinants are similarly dominant. The wave function provided by MCSCF can be written as

$$|\psi_{MCSCF}\rangle = e^{-\kappa} \sum_n \alpha_n |\{n\}\rangle, \quad (\text{I.84})$$

where $\kappa = \sum_{i,j} k_{ij} a_i^\dagger a_j$ is the anti-Hermitian operator and e^κ thus represents the rotation of the orbital basis. Both k_{ij} and the amplitudes α_n are then variably optimized.

Finally, here is what a typical classical electronic structure calculation might look like:

1. Neglecting relativistic phenomena
2. Born-Oppenheimer approximation - separation of motion of nuclei and electrons
3. Selection of basis set, selection of position of the nuclei \mathbf{R}
4. Hartree-Fock method
 - (a) Initial calculation of F_{jk} and S_{jk} elements from the equation I.74
 - (b) Initial estimate of the coefficients $c_{ki}^{(0)}$ of the matrix \mathbb{C}_0
 - (c) iterate
 - i. Calculation of the matrix \mathbb{F}_n based on the matrix \mathbb{C}_n
 - ii. Solve the Hartree-Fock Roothaan equation I.75 with \mathbb{F}_n for the new \mathbb{C}_{n+1} (several numerical methods can be used)
 - iii. If the coefficients did not converge to the desired extent, repeat iterations (c). Otherwise continue.
 - (d) Output: H-F state $|\psi_{HF}\rangle$
5. An appropriately chosen post-H-F method with the reference state $|\psi_{HF}\rangle$

Of course, not every calculation follows this axis. Sometimes relativistic effects cannot be neglected, other times BOA is not valid, or we can use MCSCF, which has an overall modified course. However, this list is sufficient for the intuition and for the following considerations involving quantum computing techniques.

I.6 Mapping to qubits

The fundamental unit of a quantum computer is a two-level quantum system called a qubit. Qubits are manipulated using single- and two-qubit gates, and individual qubits can be measured. This section will derive the possibilities for representing the many-electron problem on a quantum computer. These possibilities vary depending on whether we work in first or second quantization and whether we consider lattice methods or methods based on basis sets. However, like the classical methods described earlier, this work will only consider methods based on basis sets. The crucial difference between first and second quantization will be whether the antisymmetry is enforced directly on the wave function or by

the behavior of the operators acting on it. Simplified, it can be stated that in first quantization, the task of simulating the Hamiltonian is easier, but the task of preparing the antisymmetrized state is much more challenging. In the following, we will consider N electrons and M MOs.

I.6.1 Mapping in the first quantization

While the approach in second quantization absolutely dominates classical computational chemistry, using quantum computers reveals that the first quantization approach may have its advantages. Here, M MOs can be described using $\lceil \log_2(M) \rceil$ qubits corresponding to states $|0\dots0\rangle$ to $|1\dots1\rangle$. Each of the N electrons can then be in one or some superposition of these MOs. Overall, $N\lceil \log_2(M) \rceil$ qubits will be needed. As noted in [1], this already can be seen as a potential advantage of the first quantization approach. In second quantization, as will be shown later, M qubits are used to describe M MOs, and this number does not depend on the number of electrons. Therefore, if we have a system where it is appropriate to consider the number of possible states much larger than the number of particles, then the first quantization approach will be more qubit-efficient.

As previously mentioned, the fundamental problem with this approach is the antisymmetrization of the initial state. Considering each electron in exactly one MO and the overall state as a tensor product of these states, it is clear that this state will not be antisymmetric with respect to the exchange of electrons. It is therefore necessary to come up with an efficient algorithm for antisymmetrization. Such an algorithm was first published in [1], but a more efficient algorithm in terms of complexity was presented in [10]. The latter will be very briefly summarized here.

The goal is to perform a transformation on a register called **target**

$$|r_{N-1}\dots r_0\rangle \longrightarrow \sum_{\sigma \in S_N} (-1)^{\pi(\sigma)} |\sigma(r_{N-1}, \dots, r_0)\rangle, \quad (\text{I.85})$$

where $\pi(\sigma)$ is the sign of the permutation σ and where we require the ordering of the initial state $r_p < r_{p+1}$. The procedure can be summarized in 4 steps:

1. Preparation of the **seed** register. For a function f such that $f(N) \geq N^2$, we prepare an ancilla register **seed** in an equilibrium superposition of N -qubit states numbered 0 to $f(N) - 1$. If $f(N)$ is a power of two, then this is done through a simple Hadamard transformation.
2. Sorting of the **seed** register using a sorting network. This operation must of course be done reversibly, which requires the use of another ancilla register called **record**, where the actions of individual comparators in the network are recorded.
3. Elimination of strings with repeated entries from the **seed** register. These arise as a result of sorting the original equilibrium superposition. However, since fermionic antisymmetry is required, these repetitions must be eliminated. Detection of whether repetitions are present is done by measuring on the **seed** register. Specifically, choosing $f(N) \geq N^2$ ensures a probability of success greater than 1/2, i.e., that repetitions will not be present. Moreover, the resulting state on the **seed** register is not in any way linked with the **record** register and thus does not need to be further considered.
4. Application of reverse sorting to the **target** register. Using the values of comparators stored on the **record** register, each sorting step can be applied to the **target** register. The resulting state on this register is then a balanced superposition of all permutations of the original values. Antisymmetrization is ensured by the application of a controlled phase gate after each swap.

As stated in the article, the total complexity of the circuit is then $O(N \log_2^c(N) \log_2(M))$, where $c \geq 1$ is a constant depending on the choice of the sorting network and the algorithm uses $O(N \log_2(N))$ ancilla qubits.

The Hamiltonian is obtained by projecting it onto the one-particle basis functions as

$$H = \sum_{i=0}^{N-1} \sum_{p,q=0}^{M-1} h_{pq} |\phi_q\rangle_i \langle \phi_p|_i + \frac{1}{2} \sum_{i,j=0, i \neq j}^{N-1} \sum_{p,q,r,s=0}^{M-1} h_{pqrs} |\phi_p\rangle_i |\phi_q\rangle_j \langle \phi_r|_j \langle \phi_s|_i, \quad (\text{I.86})$$

where h_{pq} and h_{pqrs} are given entirely analogously to equations I.27 and I.28. As mentioned, each MO is assigned a string of $\lceil \log_2(M) \rceil$ qubits. For example, the first sum of equation I.86 can be decomposed into a sum of tensor products, whose individual terms are of type $|x\rangle\langle x|$, $x \in 0, 1$. These can easily be expressed using Pauli matrices. The Hamiltonian can thus contain up to $O(N^2 M^6)$ terms formed by Pauli matrices. Under such obtained Hamiltonian, we can allow the antisymmetrized wave function to evolve and thereby obtain (as will be shown later) the ground state

I.6.2 Mapping in second quantization

The second quantization formalism seems to be a suitable and very natural way to represent fermionic systems on a quantum computer - the occupation of each state is either 0 or 1, which can be automatically mapped to qubits (though it depends on the chosen mapping scheme). From this, it is directly apparent that M qubits will be needed for the computation. On the other hand, here we need to map fermionic creation and annihilation operators (acting on indistinguishable fermions) to operators acting on distinguishable qubits. There are three basic mapping schemes (and many derived from them) differing both in the degree of compromise between the non-locality of parity and occupation operators, and in the complexity of the mapping itself. First introduced is the Jordan-Wigner mapping, which is the most intuitive. Then, the parity mapping will be mentioned, which serves more as a good explanatory tool than having significant application itself. Finally, the Bravyi-Kitaev mapping will be discussed, which is a compromise between the previously mentioned and offers the best complexity.

I.6.2.1 Jordan-Wigner mapping

This mapping [29] actually uses the most intuitive approach, where a qubit in the $|0\rangle$ state corresponds to the MO occupation number 0 and the $|1\rangle$ state to the occupation number 1. If we want to represent the creation and annihilation operators on the Hilbert space of qubits so that the relations I.13 and I.15 are satisfied, then obviously it suffices to define

$$a_p^\dagger = \mathbb{I}_{M-1} \otimes \cdots \otimes \mathbb{I}_{p+1} \otimes \sigma_p^+ \otimes Z_{p-1} \otimes \cdots \otimes Z_0, \quad (\text{I.87})$$

$$a_p = \mathbb{I}_{M-1} \otimes \cdots \otimes \mathbb{I}_{p+1} \otimes \sigma_p^- \otimes Z_{p-1} \otimes \cdots \otimes Z_0, \quad (\text{I.88})$$

where $\sigma^+ = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = |1\rangle\langle 0| = \frac{1}{2}(X - iY) = (\sigma^-)^\dagger$ and X, Y, Z are the standard Pauli matrices. It is seen that the chain of Pauli Z matrices calculates the required phase $(-1)^{\sum_{i=0}^{p-1} n_i}$. If we define the Pauli weight as the number of non-identity Pauli matrices in chains like I.87, we notice that the particle number operator $a_p^\dagger a_p$ will be local, whereas other operators like $a_p^\dagger a_q$ can be very non-local due to the large Pauli weight of parity. Generally, the chain representing creation or annihilation operators has a Pauli weight of $O(M)$. It's worth mentioning that the Jordan-Wigner basis beautifully illustrates the advantage of quantum computers. While on a classical computer, the memory required to store the FCI wave function

scaled exponentially with the number of electrons (roughly as $O(M^N)$), on a quantum computer, only M qubits are needed.

In the Jordan-Wigner mapping, therefore, the transformation from fermions to qubits

$$|n_{M-1} \dots n_0\rangle \longrightarrow |q_{M-1} \dots q_0\rangle, q_i = n_i \in \{0, 1\} \quad (\text{I.89})$$

stores occupation locally and parity non-locally. The exact opposite approach is chosen in the so-called parity mapping, where the state of individual qubits does not reflect the occupation of a given MO, but the parity of the occupation of all previous MOs and it itself. The transformation of states thus takes the form

$$|n_{M-1} \dots n_0\rangle \longrightarrow |q_{M-1} \dots q_0\rangle, q_i = \left(\sum_{k=0}^i n_k \right) \pmod{2}. \quad (\text{I.90})$$

Hence, here the parity is localized, which means that when mapping the creation and annihilation operators, the parity operators will have a small Pauli weight, and conversely, the occupation operators will have a large weight. Since the parity information stored in all the following qubits must change when the occupation of some MO is changed, it is not difficult to derive the operator mapping:

$$a_p^\dagger = X_{M-1} \otimes \dots \otimes X_{p+1} \otimes P_{p,p-1}^+ \otimes \mathbb{I}_{p-2} \otimes \dots \otimes \mathbb{I}_0, \quad (\text{I.91})$$

$$a_p = X_{M-1} \otimes \dots \otimes X_{p+1} \otimes P_{p,p-1}^- \otimes \mathbb{I}_{p-2} \otimes \dots \otimes \mathbb{I}_0, \quad (\text{I.92})$$

where $P_{p,p-1}^\pm$ are two-qubit operators defined as

$$P_{p,p-1}^\pm = \sigma_p^\pm \otimes |0\rangle\langle 0|_{p-1} - \sigma_p^\mp \otimes |1\rangle\langle 1|_{p-1} = \frac{1}{2}(X_p \otimes Z_{p-1} \mp iY_p \otimes \mathbb{I}_{p-1}). \quad (\text{I.93})$$

The necessity of using such operators arises from the fact that if we want to simulate the creation (or annihilation) of an electron in orbital p in this mapping, it is necessary to apply σ_p^+ or σ_p^- depending on the state of qubit $(p-1)$. It is easily seen that if qubit $(p-1)$ is in the state $|0\rangle$, then the state of qubit p corresponds to the occupancy of orbital p , and it suffices to apply σ_p^+ in the case of a creation operator and σ_p^- in the case of an annihilation operator. However, if qubit $(p-1)$ is in the state $|1\rangle$, then the parity of qubit p will be exactly opposite to the occupancy of orbital p , and it is therefore necessary to act with σ_p^- in the case of a creation operator and σ_p^+ in the case of an annihilation operator.

Clearly, the transformation between the Jordan-Wigner and parity bases is given as

$$|q_{M-1}\rangle \otimes \dots \otimes |q_0\rangle \longrightarrow |q_{M-1} \oplus \dots \oplus q_0\rangle \otimes \dots \otimes |q_1 \oplus q_0\rangle \otimes |q_0\rangle, \quad (\text{I.94})$$

where \oplus means addition modulo 2. If we consider states as vectors (i.e. state $|n_{M-1} \dots n_0\rangle$ equivalent to vector $(n_{M-1}, \dots, n_0)^T$), then it is not difficult to estimate the form of the matrix of this transformation, denoted as \mathbb{P}

$$[\mathbb{P}]_{ij} = \begin{cases} 1 & \iff i < j \\ 0 & \iff i \geq j \end{cases}, \quad \text{hence } \mathbb{P} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ 0 & 1 & \dots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{pmatrix}, \quad \text{which gives } q_i^{\text{parity}} = \sum_{j=0}^{M-1} [\mathbb{P}]_{ij} q_j^{\text{JW}} \pmod{2}. \quad (\text{I.95})$$

The problem with the parity basis is that, compared to Jordan-Wigner, it does not reduce the necessary number of used Pauli operators to simulate creation or annihilation operators, the Pauli weight remains $O(M)$ in general. This problem is improved by the third mapping considered.

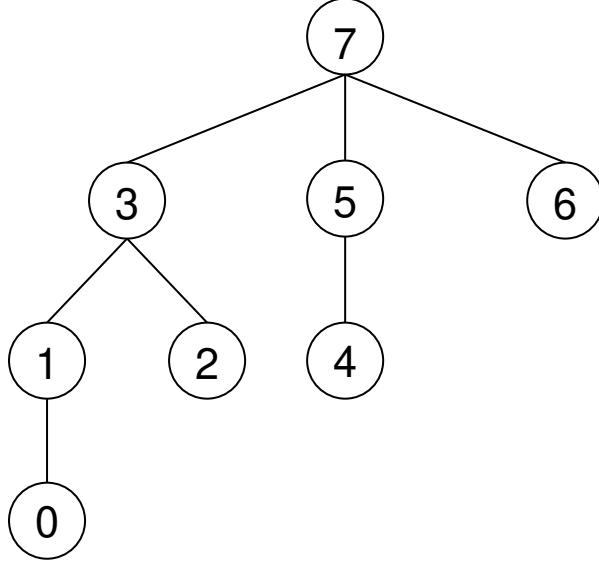


Figure I.1: A binary tree structure reflecting the Bravyi-Kitaev transformation for $M = 8$ MOs. The structure shows how each qubit holds information about the parity of its own and all qubits in the branches below it, and thus represents the equation I.97.

I.6.2.2 Bravyi-Kitaev mapping

This mapping falls somewhere between the previously mentioned ones in that it balances between the locality of storing information about occupancy and parity. It is notable for reducing the Pauli weight required to simulate creation and annihilation operators to $O(\log_2(M))$, which is a highly desired improvement for current NISQ computers. The basic idea of this mapping is that each qubit stores partial sums of occupancy numbers according to the algorithm explained below.

Generally, it again involves a transformation

$$|n_{M-1} \dots n_0\rangle \longrightarrow |q_{M-1} \dots q_0\rangle, \quad (\text{I.96})$$

where we define q_j below. The easiest way to approach this encoding is to start with an example for $M = 8$ MOs. In such a case, the value of the qubits will be given as

$$\begin{aligned} q_7 &= n_0 + n_1 + n_2 + n_3 + n_4 + n_5 + n_6 + n_7 \\ q_3 &= n_0 + n_1 + n_2 + n_3 \\ q_1 &= n_0 + n_1 & q_5 &= n_4 + n_5 \\ q_0 &= n_0 & q_2 &= n_2 & q_4 &= n_4 & q_6 &= n_6. \end{aligned} \quad (\text{I.97})$$

In it, the structure of a binary tree can be seen, see fig. (I.1) (in particular, it can be noticed that qubits with an even index hold information only about themselves, while qubits with an index $2^k - 1, k \in \mathbb{N}$, holds information about the parity of all qubits below it). In order to correctly express the value of q_i mathematically, we will represent the qubit indices in a natural way using a binary string. However, we need all strings to be the same length, so they will be padded with zeros if needed, i.e. $3 = 11 = 011$ for example. We introduce a partial ordering \leq on these strings as follows. We say that $\alpha_{l-1} \dots \alpha_0 \leq \beta_{l-1} \dots \beta_0$ if $\alpha_l = \beta_l$ for $l \geq l_0$ and at the same time $\beta_{l_0-1} = \dots = \beta_0 = 1$ for some l_0 . Like this, for

example

$$111 > \begin{cases} 011 > \begin{cases} 001 > 000 \\ 010 \end{cases} \\ 101 > 100 \\ 110 \end{cases}, \quad (\text{I.98})$$

where $j > k$ just means $j \geq k, j \neq k$. It can be noted that if $j < k$, then $j < k$. With the arrangement defined in this way, we can already define the value of q_j for any M as

$$q_j = \sum_{k \leq j} n_k (\mod 2). \quad (\text{I.99})$$

The already apparent logarithmic complexity becomes clearer as the parity of each qubit effectively manifests on every qubit with index $2^k - 1$ greater than the considered qubit, i.e., the next necessary update of qubits with a higher index occurs only when their number doubles (this applies except for a constant number of closely followed qubits, which also contain parity information and thus need to be updated). This becomes more evident if we again identify in the Jordan-Wigner basis the state $|n_{M-1} \dots n_0\rangle$ with the vector $(n_{M-1}, \dots, n_0)^T$. The transformation to the Bravyi-Kitaev basis is mediated by the matrix β , thus

$$q_j = \sum_{k=0}^{M-1} [\beta]_{jk} n_k (\mod 2). \quad (\text{I.100})$$

The form of the matrix β can be obtained from the expression I.99 and is defined recursively. Let β_n denote the transformation matrix from the Jordan-Wigner and Bravyi-Kitaev basis acting on the basis vector of length n . Obviously $\beta_1 = (\beta_1)^{-1} = (1)$. Based on this, we then define

$$\beta_{2^{k+1}} = \begin{pmatrix} \beta_{2^k} & \mathbb{A} \\ \mathbb{O} & \beta_{2^k} \end{pmatrix} \quad \beta_{2^{k+1}}^{-1} = \begin{pmatrix} \beta_{2^k}^{-1} & \mathbb{O}' \\ \mathbb{O} & \beta_{2^k}^{-1} \end{pmatrix}, \quad (\text{I.101})$$

where \mathbb{O} is a $(2^k \times 2^k)$ zero matrix, \mathbb{O}' is a $(2^k \times 2^k)$ zero matrix with a one at position $[0, 0]$ (numbering from 0 from the top left, i.e., the opposite of how qubits are indexed), and \mathbb{A} is a $(2^k \times 2^k)$ zero matrix with ones across the entire first (indexed as zeroth) row. It's worth noting that if $2^k < n < 2^{k+1}$ (i.e., n is not a power of two), then β_n is just the $(n \times n)$ segment of the matrix $\beta_{2^{k+1}}$ acting on qubits indexed from 0 to $n - 1$ (i.e., the bottom right square matrix). However, it's often the case that the number of MOs is chosen to be a power of two. Now, applying such a matrix to a generic vector $(n_{M-1}, \dots, n_0)^T$ begins to reveal the structure of the logarithmic occurrence of information about the occupation of individual MOs. Therefore, it's not necessary (unlike with parity mapping) to adjust all qubits with an index greater than j when changing qubit q_j , but only those in whose partial sums qubit q_j appears, with the number of such sums growing logarithmically with the number of MOs.

The representation of creation and annihilation operators in the Bravyi-Kitaev basis is significantly more complex than in previous encodings and cannot be fully derived here. The complete derivation can be found in [55]. Here, let's just mention that for any index j , qubits in the Bravyi-Kitaev basis are divided into three sets. The *parity* set ($P(j)$) includes all qubits that store the parity of all MOs with an index less than j . The *update* set ($U(j)$) consists of qubits that store a partial sum including the MO with index j . Finally, the *flip* set ($F(j)$) contains qubits that determine whether the qubit with index j has the same parity as the MO with index j . We introduce the notation where for some set of indices S , O_S signifies an operator (usually some Pauli matrix) acting on qubits with indices contained in S . Finally,

sets are defined as

$$R(j) \equiv P(j) \setminus F(j)$$

$$\rho(j) \equiv \begin{cases} P(j) & \text{if } j \text{ is even} \\ R(j) & \text{if } j \text{ is odd.} \end{cases} \quad (\text{I.102})$$

In this way, the creation and annihilation operator acting on the MO with index j in the Bravyi-Kitaev basis can be represented as

$$a_j^\dagger = \frac{1}{2}[X_{U(j)} \otimes X_j \otimes (Z_{P(j)} - iX_{U(j)}) \otimes Y_j \otimes Z_{\rho(j)}],$$

$$a_j = \frac{1}{2}[X_{U(j)} \otimes X_j \otimes (Z_{P(j)} + iX_{U(j)}) \otimes Y_j \otimes Z_{\rho(j)}]. \quad (\text{I.103})$$

The products of these operators then form the Hamiltonian, with different types of the expression $a_i^\dagger a_j$ (according to the evenness or oddness of indices and the overlap of their parity and update sets) allowing for algebraic rules for representation in the Bravyi-Kitaev basis to be found, from which two-particle terms can subsequently be assembled. A detailed description of this process can also be found in [55].

Worth mentioning here is also a modified version of this algorithm called BK-tree [26]. Although on one hand, it creates operators with generally greater Pauli weight, on the other hand, this mapping is unambiguous (unlike the basic one) even for a number of MOs that is not a power of two, which subsequently allows for a reduction in the number of qubits by utilizing symmetries.

I.7 Determining the electronic structure using the phase estimation algorithm

Finding the ground state using the phase estimation algorithm (QPE) can be considered a solution intended for fault-tolerant quantum computers, due to the required circuit depth and the resulting long coherence times. Thus, it cannot be assumed that this approach would be advantageously used for these purposes in the near future. QPE is a well-known subalgorithm used within many quantum algorithms and is detailed, for example, in [42]. Instead of detailing the algorithm itself, this section will summarize its application to obtaining the ground state.

The time evolution of the initial state $|\psi(0)\rangle$ under the electronic Hamiltonian H can be written in the energy basis as

$$|\psi(\tau)\rangle = e^{-\frac{i}{\hbar}H\tau} |\psi(0)\rangle = \sum_i e^{-\frac{i}{\hbar}E_i\tau} \langle E_i|\psi(0)\rangle |E_i\rangle = \sum_i c_i e^{-i2\pi\varphi_i} |E_i\rangle, \quad (\text{I.104})$$

where $c_j \equiv \langle E_i|\psi(0)\rangle$ and phase $\varphi_i \equiv E_i\tau/(2\pi\hbar) \in [0, 1]$. In the language of binary fractions, and taking into account the round-off error arising from the finite qubit representation, the phase corresponding to the i th level can be written as

$$\varphi_i = \sum_{j=0}^{t-1} \frac{\varphi_{j,i}^{\text{bin}}}{2^j} = 0.\varphi_{1,i}^{\text{bin}} \dots \varphi_{t,i}^{\text{bin}}. \quad (\text{I.105})$$

t therefore indicates the precision of how many bits the phase will be determined on and thus the required number of ancilla qubits (register 1). Let us further denote the rounding error as ϵ_{PE} . In addition, t is also determined by the required probability of success of the algorithm. QPE can then be summarized in the following steps:

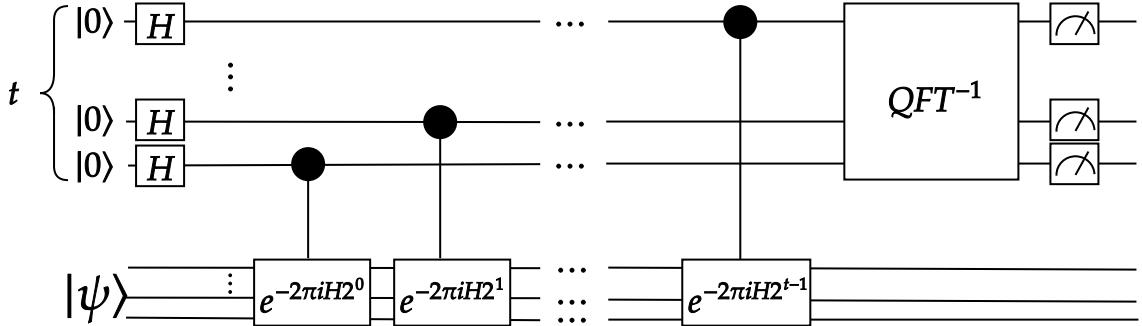


Figure I.2: QPE to determine the eigenvalues and eigenstates of the electron Hamiltonian.

1. Prepare register 2 in the $|\psi(0)\rangle$ state, which has a non-zero (and ideally as large as possible) overlap with the ground state (or another required from the spectrum of the Hamiltonian). Prepare register 1 in state $|0\rangle^{\otimes t}$.
2. Perform the Hadamard transformation on register 1 and subsequently controlled operations according to Fig. I.2 i.e.

$$\frac{1}{\sqrt{2^t}} \sum_x |x\rangle |\psi(0)\rangle \longrightarrow \frac{1}{\sqrt{2^t}} \sum_x \sum_i c_i e^{-i2\pi\varphi_i x} |x\rangle |E_i\rangle \quad (\text{I.106})$$

3. Perform an inverse Fourier transform on register 1 and thereby obtain a binary expression for φ_i

$$\frac{1}{\sqrt{2^t}} \sum_x \sum_i c_i e^{-i2\pi\varphi_i x} |x\rangle |E_i\rangle \longrightarrow \sum_i c_i |\varphi_i^{\text{bin}}\rangle |E_i\rangle. \quad (\text{I.107})$$

4. By measuring the first register in the computational base, we obtain with probability $|c_i|^2$ the phase φ_i and from this easily estimate the eigenvalue of the energy. At the same time, register 2 collapses into its own state $|E_i\rangle$.

Many variants of this algorithm have been invented, some of which can be significantly more efficient than this basic variant. For calculations in quantum chemistry, for example, classically attainable knowledge of the gap between the ground and the first excited state can be advantageously used [10].

From the procedure so far, it is clear that two things must be fulfilled to successfully implement QPE: prepare an initial state with a non-zero overlap with the target eigenstate and effectively implement the evolution operator e^{-iH} (instead of the evolution operator, one can sometimes choose an effectively invertible function of the Hamiltonian). These will be discussed in more detail below.

I.7.1 Initial state preparation

As it was mentioned, the overlap between the initial state and the desired final eigenstate of the Hamiltonian (FCI) must be nonzero, and the extent of this overlap determines the probability of measuring the desired state. A common choice for the initial state is the state obtained by the H-F method, which typically has a large overlap with the ground state. For highly correlated systems, it is advisable to use more advanced classical computational methods such as CI or MCSCF. An interesting option for state preparation is the use of variational algorithms [70], which, as will be shown in the next chapter, can serve to determine the ground state on their own. Also worth mentioning is the adiabatic state preparation

[3], where for any Hamiltonian H_s , we can obtain a state close to its ground state by preparing the ground state of some simple Hamiltonian H_0 and subsequently letting the state evolve under a Hamiltonian that slowly changes from H_0 to H_s . From the adiabatic theorem, it then follows that if certain conditions are met, the system will remain in the same instantaneous subspace throughout the process. It is common to choose H_0 such that its ground state is the state provided by the H-F method.

It is important to note that preparing a state with sufficient overlap with the desired state is absolutely crucial, not least because for a randomly selected initial state, the probability of collapsing into the desired state exponentially decreases with the size of the system.

I.7.2 Hamiltonian simulation

In this subsection, the most commonly used methods for simulating the evolution under the Hamiltonian H in quantum computational chemistry will be discussed. Thus, it will be about ways of approximating the evolution operator e^{-iHt} so that the result can be efficiently implemented on a quantum computer. It is necessary to mention that for a completely general Hamiltonian, the number of used elementary gates grows exponentially with the number of qubits [42]. Therefore, it is always necessary to assume some special structure of the Hamiltonian.

I.7.2.1 Trotterization

Trotterization is a shorthand term for the use of the Lie-Trotter-Suzuki (also product) formula stemming from [63]. The idea behind this method is quite simple. Consider a Hamiltonian H that can be expressed as $H = \sum_{i=1}^m h_i$, where h_i are local Hamiltonians that are easy to implement, and m indicates the number of terms in the Hamiltonian. It is easy to realize that the electronic Hamiltonian I.26 meets this criterion. This fermionic Hamiltonian has $O(M^4)$ terms. After mapping to qubits, it contains $O(M^4)$ Pauli strings, whose locality grows as $O(M)$ in the case of Jordan-Wigner and parity mappings and as $O(\log(M))$ in the case of Bravyi-Kitaev mapping. Exponentials of these strings are then easily implementable (details in [69]). For times $t \ll 1$, the product formula says

$$\exp\left\{-i \sum_{i=1}^m h_i t\right\} = \prod_{i=1}^m \exp\{-ih_i t\} + O(m^2 t^2). \quad (\text{I.108})$$

The error here naturally arises from the fact that the individual h_i do not commute with each other, and for an exact expansion, the BCH lemma would be needed. However, it is apparent that for a sufficiently small time interval, the error can be made arbitrarily small. This does not help, however, if the required evolutionary times are large (for example, in QPE in its simplest version, it is evident that the required evolutionary time is given by the highest power in controlled evolution).

The real significance of the product formula, however, lies in the possibility of dividing the time evolution into short segments - so-called Trotter steps. The product formula in the first order then will have the form

$$\exp\left\{-i \sum_{i=1}^m h_i t\right\} = \left(\prod_{i=1}^m \exp\left\{\frac{-ih_i t}{S}\right\} \right)^S + O\left(\frac{m^2 t^2}{S}\right). \quad (\text{I.109})$$

From this, several things can be observed. The first is that the total complexity of the simulation (meaning the number of applications of individual e^{-ih_i}) will be $O(mS)$. The second is that in the limit $S \rightarrow +\infty$, the error goes to 0. This leads to the third observation, that to achieve accuracy ϵ , it is necessary to choose $S = O(m^2 t^2 / \epsilon)$. However, it must be realized that this also increases the simulation time. A schematic drawing of the evolution simulation using the first-order product formula is shown in figure I.3.

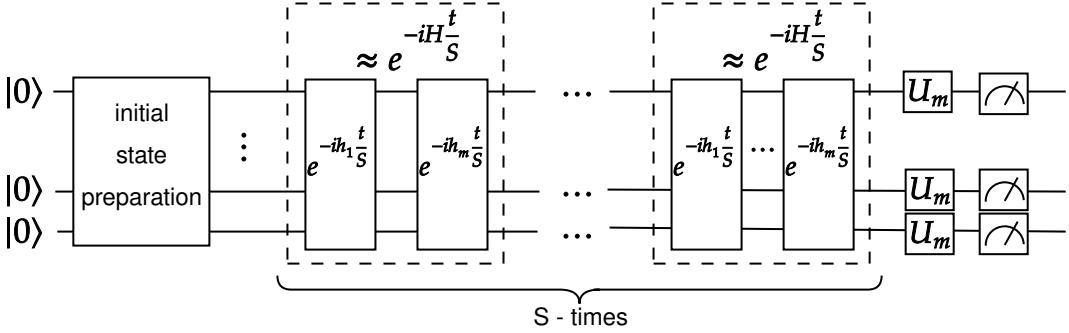


Figure I.3: Simulation of evolution under the H Hamiltonian using Trotterization. First, a state is prepared in the desired initial state (for QPE in a state with the largest possible overlap with the desired eigenstate of the Hamiltonian - e.g. H-F state for the ground state). Subsequently, Trotter steps are applied. Finally, appropriate operations are applied depending on the property being measured.

It is evident that by suitably arranging the exponentials in expression I.109, it is possible to reduce the error. Intuitively, one can imagine that if the commuting terms are first grouped together in the sum $H = \sum_{i=1}^m h_i$, then when these terms are separated into individual exponentials, no error occurs. Methods based on this idea include, for example, randomization [17].

In the mentioned work, randomization is used in conjunction with another way to reduce error - the use of higher-order product formulas. Indeed, it was Suzuki's merit to have expanded the product formula recursively to even higher orders. Here, the second order will be discussed. The second-order product formula is based on the idea that errors from the first order cancel out in a first approximation when the exponentials are reapplied in the reverse order. The evolutionary operator in this choice will be approximated as

$$\exp\left\{-i \sum_{i=1}^m h_i t\right\} = \underbrace{\left(\prod_{i=1}^m \exp\left\{\frac{-ih_i t}{2S}\right\}\right)}_{U_2(t)} \prod_{i=m}^1 \exp\left\{\frac{-ih_i t}{2S}\right\}^S + O\left(\frac{m^3 t^3}{S^2}\right). \quad (\text{I.110})$$

This time it can be seen that $S = \sqrt{m^3 t^3 / \epsilon}$ should be chosen to achieve ϵ accuracy. The recursion giving higher even orders of this formula is then given as

$$U_{2k}(t) = [U_{2k-2}(p_k t)]^2 U_{2k-2}([1 - 4p_k]t) [U_{2k-2}(p_k t)]^2 = \exp\left\{-i \sum_{i=1}^m h_i t\right\} + O\left(\frac{(mt)^{2k+1}}{r^{2k}}\right), \quad (\text{I.111})$$

where $p_k = 1/(4 - 4^{1/(2k-1)})$ [60]. In principle, such approximations of arbitrarily high orders can be constructed, but from the fourth order onwards, this construction procedure ceases to be advantageous in most practical cases. Thus, Trotterization to the second order can normally be encountered.

I.7.2.2 Qubitization

Qubitization is an alternative process for simulating a Hamiltonian using ideas from simulations via quantum walks. It was first introduced in [35] and subsequently developed in [34]. In both cases, qubitization was accompanied by subsequent *quantum signal processing*. Qubitization as such is a very general approach based on the use of two oracles, through which the Hamiltonian $H = (\langle G|_a \otimes I_s)U(|G\rangle_a \otimes I_s)$ is written as a projection of the unitary oracle U onto the state $|G\rangle_a$ created by another unitary oracle

(so-called *standard* or *block encoding*). The crucial step is then to create from U an operator W (often called *walk* operator), which anchors H in the invariant $SU(2)$ subspaces corresponding to its eigenvalues. Here, the connection with Grover's algorithm will be evident, where such anchoring occurs in the invariant subspace of Grover's iterations. W is then isomorphic to the operator $e^{-iY \otimes \cos^{-1}(H)}$ in individual spaces. Quantum signal processing is then a very general and useful way to convert $H \rightarrow f[H]$, where $f : [-1, 1] \rightarrow D$, and D represents a circle bounded by the unit circle in the complex plane. In the original works, this is (very simplistically said) used to convert $e^{-i \pm \cos^{-1}(E_\lambda)} \rightarrow e^{-i E_\lambda t}$. However, as shown in [10] and [48], for determining static quantities (such as ground state energy), subsequent quantum signal processing is not necessary, and they can be determined directly from the original expression using a phase estimation algorithm. This option is not only faster but also eliminates possible phase ambiguity, and so this approach is chosen in this work as well.

As has been said, qubitization is a very general approach to simulating evolution and can use various Hamiltonian models for this process. In this work, attention will be focused most on the linear combination of unitary operators (LCU) model, where $H = \sum_{j=1}^d \alpha_j U_j$ (it is good to realize that this exactly corresponds to the Hamiltonian as a sum of Pauli strings) and further on the model of d -sparse Hamiltonian, which directly comes from the theory of quantum walks [16], [8] and on which this method was built. Finally, it is good to mention in this introduction that although qubitization requires a larger number of qubits compared to Trotterization, it offers optimal scaling with respect to evolution time and achieved error. Therefore, it is the preferred method for Hamiltonian simulation in general. However, even using this method, the simulation of tasks that would be unsolvable for today's classical computers remains within the regime where error correction is needed. An estimate of the number of superconducting qubits along with a detailed implementation of qubitization can be found in [4].

A rigorous description of qubitization is good to start with a definition.

Definition 1 (block encoding). Signal operator H with spectral norm $\|H\| \leq 1$ is encoded in block form if we can query the unitary oracle $U : \mathcal{H}_s \otimes \mathcal{H}_a \rightarrow \mathcal{H}_s \otimes \mathcal{H}_a$ and unitary oracle to prepare the state $G|0\rangle_a = |G\rangle_a \in \mathcal{H}_a$, which have the property $(\langle G|_a \otimes I_s)U(|G\rangle_a \otimes I_s) = H$. Furthermore, we assume that we also have the possibility to query the inversions and controlled versions of U and G .

Consider any input state of the system $|\psi\rangle_s$, on which we want to apply a transformation mediated by the operator H , which is not necessarily unitary. Since only unitary transformations can be performed on a quantum computer, it is necessary to realize H within a larger Hilbert space, where the action of H can be extracted from some unitary operation. This is exactly mediated by the operator U , which acts on the ancilla space \mathcal{H}_a in addition to \mathcal{H}_s . The action of H is then, according to the previous definition, encoded on the subspace designated as the *signal state* $|G\rangle_a \in \mathcal{H}_a$. Thus

$$U|G\rangle_a |\psi\rangle_a = |G\rangle_a H |\psi\rangle_a + \sqrt{1 - \|H|\psi\rangle_s\|^2} |G_\psi^\perp\rangle_{as}, \quad U = \begin{pmatrix} H & \cdot \\ \cdot & \cdot \end{pmatrix}, \quad (\langle G|_a \otimes I_s) |G_\psi^\perp\rangle = 0. \quad (\text{I.112})$$

Thus, the signal state $|G\rangle$ defining the measurement basis on \mathcal{H}_a divides the action of U into two subspaces. On $\mathcal{H}_G = |G\rangle \otimes \mathcal{H}_s$ the measurement succeeds with probability $\|H|\psi\rangle_s\|^2$ and $U|G\rangle_a |\psi\rangle_a$ is projected onto $(|G\rangle H|\psi\rangle)/\|H|\psi\rangle\|$. On the orthogonal complement \mathcal{H}_{G^\perp} the measurement fails. This probabilistic interpretation also results in the requirement for the spectral norm $\|H\| \leq 1$. In what follows, the unspecified parts of U that transform $|G_\psi\rangle$ into an orthogonal complement and are not important for the following derivation will be omitted.

For H Hermitian, the action of U in the basis of the eigenstates of H can be written, $H|\lambda\rangle = \lambda|\lambda\rangle$, as

$$U|G\rangle|\lambda\rangle = U|G_\lambda\rangle = \lambda|G_\lambda\rangle + \sqrt{1 - |\lambda|^2} |G_\lambda^\perp\rangle. \quad (\text{I.113})$$

Let's denote the subspace corresponding to the eigenvalue λ as $\mathcal{H}_\lambda = \text{span}\{|G_\lambda\rangle, U|G_\lambda\rangle\}$. Notice that if we now want to implement some function of the Hamiltonian $f[H]$ in block form (which is the goal of quantum signal processing) - for example, H^2 to determine variance - then a problem occurs due to the fact that \mathcal{H}_λ are not invariant subspaces for U , and its repeated application therefore moves us outside these subspaces. The solution is to find an operator W , called an *iterator*, which on one hand still satisfies I.112 and preserves the property $(\langle G_a | \otimes I_s)W(|G\rangle_a \otimes I_s) = H$, on the other hand, however, for each eigenstate of the Hamiltonian $|\lambda\rangle$, performs an SU(2) rotation in disjoint 2D subspaces $\mathcal{H}_\lambda = \text{span}\{|G_\lambda\rangle, W|G_\lambda\rangle\} = \text{span}\{|G_\lambda\rangle, |G_\lambda^\perp\rangle\}$. This defines the state $|G_\lambda^\perp\rangle$ through the Gram-Schmidt orthogonalization process as

$$|G_\lambda^\perp\rangle = \frac{(W - \lambda)|G_\lambda\rangle}{\sqrt{1 - |\lambda|^2}} \quad (\text{I.114})$$

and further the action of Pauli matrices in individual subspaces as

$$X_\lambda|G_\lambda\rangle = |G_\lambda^\perp\rangle \quad Y_\lambda|G_\lambda\rangle = i|G_\lambda^\perp\rangle \quad Z_\lambda|G_\lambda\rangle = |G_\lambda\rangle. \quad (\text{I.115})$$

In this basis, the iterator on the individual subspaces is written as

$$W_\lambda = \begin{pmatrix} \lambda & -\sqrt{1 - |\lambda|^2} \\ \sqrt{1 - |\lambda|^2} & \lambda \end{pmatrix}. \quad (\text{I.116})$$

Here it is certainly possible to identify $\lambda \equiv \cos(\theta_\lambda)$. In this way, in W_λ we already recognize the rotation around the y axis and can thus be written as

$$W = \bigoplus_\lambda \begin{pmatrix} \cos(\theta_\lambda) & -\sin(\theta_\lambda) \\ \sin(\theta_\lambda) & \cos(\theta_\lambda) \end{pmatrix}_\lambda = \bigoplus_\lambda e^{-iY_\lambda\theta_\lambda}, \quad (\text{I.117})$$

where $\theta_\lambda = \cos^{-1}(\lambda)$.

However, can such a W be created from the block form of H at all? And can it be implemented efficiently? The answer to these questions is the very essence of qubitization and is summarized in the following theorem.

Theorem 1 (qubitization). Let H be a Hermitian matrix encoded in the standard form $H = \langle G_a | U | G_a \rangle$. Then the iterator W from the equation I.117 can be constructed using at most one query of each oracle G , controlled by U , their inversions, one additional qubit, and using $O(\log(\dim(\mathcal{H}_a)))$ quantum gates.

The proof of this theorem is addressed in the entirety of section 4 of the work [34] and will therefore not be discussed in detail here. In the same work, the fact that the application of W^N effectively produces Chebyshev polynomials of the N -th order, which is further used to approximate the already mentioned function $f[H] = A[H] + iB[H]$, i.e., during quantum signal processing, is also discussed in detail. However, for this work, the expression I.117, on which the phase estimation algorithm can be directly applied, is sufficient. Next, the encoding will be demonstrated first on the LCU model and then on the model of the d -sparse Hamiltonian.

LCU model

This model, where $H = \sum_{j=1}^d \alpha_j U_j$, is probably the most suitable in terms of use in quantum chemistry due to its structure, which corresponds to the sum of Pauli strings obtained by mapping from fermions to qubits. The encoding into block form here stems from the work [18], which laid the foundation for subsequent simulation of Hamiltonian dynamics using truncated Taylor series [9].

In this model, therefore

$$H = \sum_{j=1}^d \alpha_j U_j, \quad \|H\| \leq \|\alpha\|_1 = \sum_{j=1}^d |\alpha_j|, \quad (\text{I.118})$$

where $\|\alpha\|$ is the upper bound for the spectral norm. By absorbing the complex phase into U_j , it can be chosen that $\alpha_j \geq 0$. The algorithm assumes that α_j are provided as a list of d numbers and that all U_j are implemented as a quantum circuit using $O(C)$ elementary quantum gates. Again, it is good to note that these requirements are met in the decomposition into Pauli strings considered here, where α_j correspond to orbital integrals I.27, I.28.

It can be easily verified that by choosing an oracle

$$G = \sum_{j=1}^d \sqrt{\frac{\alpha_j}{\|\alpha\|}} |j\rangle_a \langle 0|_a, \quad U = \sum_{j=1}^d |j\rangle_a \langle j|_a \otimes U_j, \quad (\text{I.119})$$

where G can be constructed using $O(d)$ elementary gates and U using $O(dC)$ elementary gates, we achieve the desired property

$$\langle G | U | G \rangle = \frac{H}{\|\alpha\|}, \quad (\text{I.120})$$

where $|G\rangle = G|0\rangle$. Indeed,

$$\langle (G|_a \otimes I_s) U (|G\rangle_a \otimes I_s) = \langle (G|_a \otimes I_s) \sum_{j=1}^d \sqrt{\frac{\alpha_j}{\|\alpha\|}} |j\rangle_a \otimes U_j = \sum_{j,k=1}^d \frac{\sqrt{\alpha_j \alpha_k}}{\|\alpha\|} \langle k|j\rangle \otimes U_j = \frac{1}{\|\alpha\|} \sum_{j=1}^d \alpha_j U_j. \quad (\text{I.121})$$

It can also be seen from this that the subsequent calculation of the spectrum will need to be rescaled by the norm $\|\alpha\|$. By relabeling $\lambda \rightarrow E_\lambda$ and rescaling, the expression I.116 takes the form

$$W_\lambda = \begin{pmatrix} E_\lambda & -\sqrt{1 - \frac{|E_\lambda|^2}{\|\alpha\|^2}} \\ \sqrt{1 - \frac{|E_\lambda|^2}{\|\alpha\|^2}} & E_\lambda \end{pmatrix}_\lambda. \quad (\text{I.122})$$

This operation can be diagonalized and from the expression I.117 it is clear that the eigenstates here will be the eigenstates of the Y_λ matrix, i.e. the states of $(|G_\lambda\rangle \pm i|G_\lambda^\perp\rangle)/\sqrt{2} \equiv |\pm G_\lambda\rangle$. Thus

$$W|\pm G_\lambda\rangle = e^{\mp i \cos^{-1}(\frac{E_\lambda}{\|\alpha\|})} |\pm G_\lambda\rangle. \quad (\text{I.123})$$

Here it is already possible to use the phase estimation algorithm, which will provide the value $\mp \cos^{-1}(\frac{E_\lambda}{\|\alpha\|})$. The eigenvalue of the energy is then easily obtained as

$$E_\lambda = \|\alpha\| \cos[\cos^{-1}(\frac{E_\lambda}{\|\alpha\|})]. \quad (\text{I.124})$$

d -sparse Hamiltonian model

In this model, the Hamiltonian has at most d nonzero elements in each row. At the same time, access to their position and H_{jk} values is mediated through the oracle [8]

$$O_H |j\rangle |k\rangle |z\rangle = |j\rangle |k\rangle |z \oplus H_{jk}\rangle \quad O_F |j\rangle |l\rangle = |j\rangle |f(j, l)\rangle. \quad (\text{I.125})$$

This model is commonly used in algorithms based on quantum walks [16], [8]. However, for the purposes of quantum chemistry, it may not be the most suitable due to the need for an efficient method of calculating the matrix elements of the Hamiltonian, which in the case of quantum chemistry can be very costly [10]. However, it is mentioned here because the theory of qubitization was developed on this model.

From I.125, it is seen that O_H takes row j and column k and returns H_{jk} in binary format. O_F takes row j and a number $l \in \{1, \dots, d\}$ and returns $f(j, l)$ representing the index of the l -th nonzero element in row j . Let $F_j = \{f(j, l)\}_l$ be the set of all column indices with nonzero elements in row j . The encoding into block form in this case is done by choosing $G = I$ (i.e., $|G\rangle = |0\rangle_{a_1} |0\rangle_{a_2} |0\rangle_{a_3} \equiv |0\rangle_a$) and $U = T_2^\dagger T_1$, where T_1, T_2 are isometries defined as

$$T_1 = \sum_j |\psi_j\rangle \langle 0|_a \langle j|_s, \quad |\psi_j\rangle = \sum_{p \in F_j} \frac{|p\rangle_{a_3}}{\sqrt{d}} \left(\sqrt{\frac{H_{pj}}{\|H\|_{\max}}} |0\rangle_{a_1} + \sqrt{1 - \frac{|H_{pj}|}{\|H\|_{\max}}} |0\rangle_{a_1} \right) |0\rangle_{a_2} |j\rangle_s \quad (\text{I.126})$$

$$T_2 = \sum_k |\chi_k\rangle \langle 0|_a \langle k|_s, \quad \langle \chi_k| = \sum_{q \in F_k} \frac{\langle q|_s}{\sqrt{d}} \left(\sqrt{\frac{H_{kq}}{\|H\|_{\max}}} \langle 0|_{a_2} + \sqrt{1 - \frac{|H_{kq}|}{\|H\|_{\max}}} \langle 0|_{a_2} \right) \langle 0|_{a_1} \langle j|_{a_3}. \quad (\text{I.127})$$

Indeed, by direct calculation it can be verified that with this choice

$$(\langle G|_a \otimes I_s) U (|G\rangle_a \otimes I_s) = (\langle 0|_a \otimes I_s) T_2^\dagger T_1 (|0\rangle_a \otimes I_s) = \sum_{k,j} \frac{H_{kj}}{d\|H\|_{\max}} |k\rangle \langle j| = \frac{H}{d\|H\|_{\max}}. \quad (\text{I.128})$$

[8] also derives how these isometries can be implemented using two queries on O_H and one query on O_F . Obtaining the energy values is then analogous to the LCU. However, in this case it is possible to state that the iterator W , which has the meaning of one step of the quantum walk, has the form

$$W = iS(2TT^\dagger - I), \quad (\text{I.129})$$

where $S |j, k\rangle = |k, j\rangle$ is the swap operator. This shape already flows directly from discrete-time quantum walks [16].

I.8 Ground state determination using VQE

As already mentioned, the circuit depth required for Quantum Phase Estimation (QPE) and the resulting necessary coherence times are (even when using the best known Hamiltonian simulation techniques to date) far beyond the capabilities of current and near-future Noisy Intermediate-Scale Quantum (NISQ) devices. A natural thought, therefore, is to try to combine the advantages of classical computers with those of quantum ones. This approach gives rise to so-called hybrid quantum-classical algorithms. The most famous representative of this class and reliably the most used algorithm for experimental computation in quantum chemistry is the *Variational Quantum Eigensolver* (VQE) [46] [40] [50].

VQE essentially represents nothing other than the use of the time-independent variational principle. This approach replaces the need for long coherence times with the ability on a quantum computer to repeatedly prepare and measure a parameterized wave function. Moreover, it provides great flexibility in the form of choosing the circuit ansatz for preparing the parameterized state, thus allowing for consideration of the specifics of the given hardware.

The Rayleigh-Ritz variational principle states that for a parameterized wave function $|\psi(\theta)\rangle$, it holds that

$$\langle H \rangle(\theta) = \langle \psi(\theta) | H | \psi(\theta) \rangle \geq E_0, \quad (\text{I.130})$$

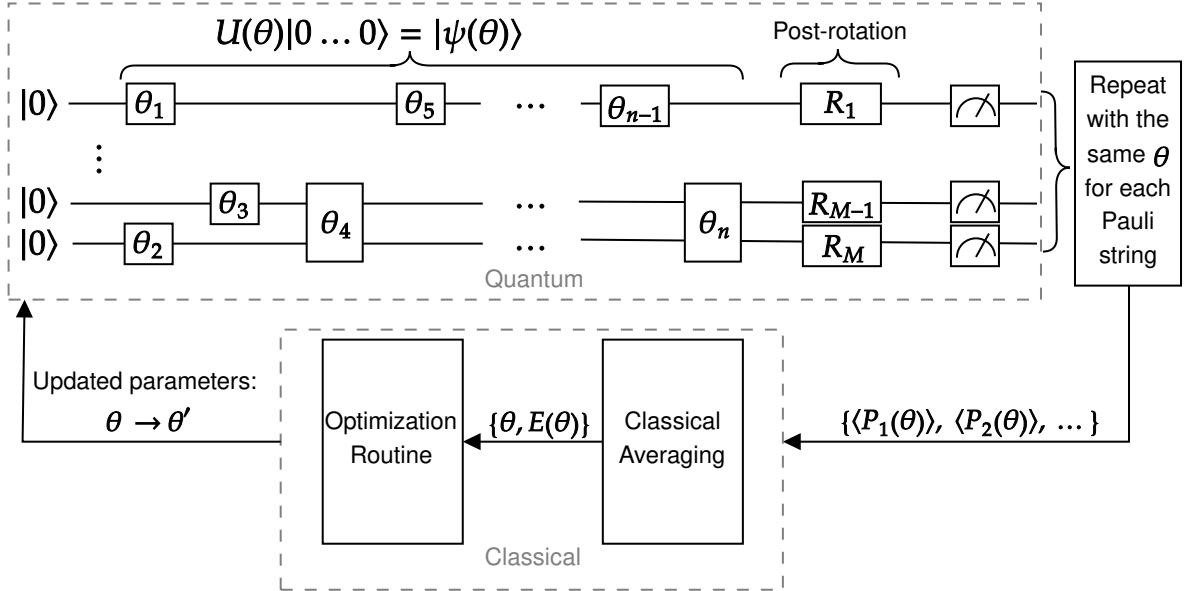


Figure I.4: Illustration of the VQE algorithm. A quantum computer prepares a parameterized state $|\psi(\theta)\rangle$. This is followed by a rotation given by which term of the Hamiltonian is currently being measured. The quantum part finishes with the measurement. A classical computer averages the outputs $\langle P_i(\theta)\rangle$ of the individual terms of the Hamiltonian, thereby obtaining the value $E(\theta)$. This represents the value of the *cost function* (cost function), which is subsequently used in the classical optimization process, the output of which is a new set of parameters θ' . These are used to prepare the $|\psi(\theta')\rangle$ state, which will ideally provide a lower energy. The whole process is repeated until the energy converges.

where H is the Hamiltonian of the system, E_0 the energy of the ground state, and $\theta = (\theta_1, \dots, \theta_n)^T$ the vector of independent parameters. Very roughly then, VQE can be divided into three steps that are iterated:

1. On a quantum computer, prepare the state $|\psi(\theta)\rangle = U(\theta)|\psi_0\rangle$, where $|\psi_0\rangle$ is easy to prepare state (e.g. computational basis) and $U(\theta)$ is given by the choice of ansatz.
2. Measure the expected energy $\langle H \rangle(\theta)$ using the *Hamiltonian averaging* procedure.
3. Use an appropriate classical optimization process to determine new values of θ that will reduce the value of $\langle H \rangle(\theta)$.
4. Repeat these steps until the energy value converges. The value of the parameters θ during convergence then defines the ground state.

This is shown schematically in the figure I.4. Next, the individual steps of VQE will be analyzed in detail.

I.8.1 Choice of ansatz for state preparation

As already mentioned, the ability to choose how the parameterized state approximating the ground state is created ensures the flexibility so necessary for NISQ devices that VQE provides. When choosing an ansatz, two things must be considered. The first is the capabilities of the device on which the computation will be carried out. This includes the available number of qubits, their connectivity, and

many other factors that need to be considered. The second is the structure of the parameterized state produced by the ansatz. This proves to be extremely important for subsequent classical optimization. It has been shown [39] that with a random choice of initial parameters, the explored Hilbert space is predominantly formed by so-called *barren plateaus* - regions with almost zero gradient in all directions, making classical optimization extremely challenging. This problem moreover worsens exponentially with the number of qubits, signaling that solutions based purely on hardware considerations will not be scalable. Intuitively, however, one might suspect (and the reality of experiments suggests) that the right choice of ansatz structure, which closely matches the physical nature of the problem, will provide measurable gradients, thus better guiding classical optimization and making it much more efficient.

In light of the previous paragraph, the design of the circuit ansatz can roughly be divided into two types:

1. Physically (chemically) motivated ansatz - circuits inspired by classical numerical methods to systematically approximate the wave function
2. Hardware heuristic ansatz - the circuits here are composed of single-qubit rotations and connecting blocks designed to make the best possible use of the capabilities of the specific hardware.

Although several physically motivated ansatz have been presented [19], [51], [56], [67], only one representative of each type will be discussed in detail in this work.

I.8.1.1 Unitary Coupled Cluster

Since this physically motivated method is directly based on its classical predecessor, the Coupled Cluster (CC) theory (already mentioned in chapter I.5), this method will be summarized here in a little more detail. It is worth recalling that here (as in Configuration Interaction (CI)) the excitation operator has the form

$$T = \sum_i T_i, \quad (\text{I.131})$$

where

$$T_1 = \sum_{i \in ex, \alpha \in occ} t_{i\alpha} a_i^\dagger a_\alpha \quad T_2 = \sum_{i,j \in ex, \alpha, \beta \in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta \quad T_3 = \dots \quad (\text{I.132})$$

The Ansatz for this method is then given as

$$|\psi_{CC}\rangle = e^T |\psi_{HF}\rangle. \quad (\text{I.133})$$

To make the method computationally feasible, mostly only one-electron and two-electron excitations are considered. The method using this shortened excitation operator is then called Coupled Cluster Single and Double (CCSD). The goal is then to determine the t coefficients in the expressions I.132 and, as a result, of course, the energy of the ground state. This can be done in the following way. The Schrödinger equation has the form

$$H |\psi_{CC}\rangle = H e^T |\psi_{HF}\rangle = E e^T |\psi_{HF}\rangle. \quad (\text{I.134})$$

Let $\{|\mu\rangle\}$ denote the set of Slater determinants corresponding to excited states up to order m (i.e., if only one- and two-electron excitations are considered, then this set consists of all possible once and twice excited Slater determinants). Multiplying e^{-T} from the left and projecting onto this set results in a set of equations

$$\begin{aligned} \langle \psi_{HF} | e^{-T} H e^T | \psi_{HF} \rangle &= E \\ \langle \mu | e^{-T} H e^T | \psi_{HF} \rangle &= 0, \end{aligned} \quad (\text{I.135})$$

where the lower set of equations is used to determine the amplitudes t and the upper equation determines the energy.

It is worth reminding that the main advantage of the Coupled Cluster (CC) method over Configuration Interaction (CI) was its size extensivity. A method that is size extensive for a system of non-interacting fragments has a wave function that is multiplicatively separable and energy that is proportionally related to the size of the system. In other words, the overall wave function can be factorized into the product of individual fragments and the total energy will be the sum of the energies of these fragments [51]. This property is very much desired when making approximate calculations of some chemical properties. The fact that this is achieved in the case of CC can be shown from the properties of the similarity-transformed Hamiltonian $\tilde{H} = e^{-T} H e^T$ [27].

The fact that equations I.135 are solvable without further approximations follows from the BCH formula

$$\tilde{H} = e^{-T} H e^T = H + [H, T] + \frac{1}{2} [[H, T], T] + \frac{1}{3!} [[[H, T], T], T] + \frac{1}{4!} [[[[H, T], T], T], T], \quad (\text{I.136})$$

which, due to the anticommutation properties of fermionic operators when applied to a single Slater determinant (which the Hartree-Fock (H-F) state obviously is), ends at the fourth order [7].

Therefore, the truncated CC method is classically solvable and provides very good results for many systems. It has thus become something of a gold standard in current classical computational chemistry. However, it has two significant drawbacks. The first is that the finiteness of the expansion I.136 only applies when applied to a single Slater determinant. This results in weak results for strongly correlated systems - for example, for transition states and generally geometries where the excited PES is nearly degenerate with the ground state. Although methods have been developed for application to a multireference initial state, the resulting complexity of $O(M^7)$ limits the use of CC to small molecules only. The second is that the operator e^T is not unitary, which prevents approaching the equation for energy I.135 variationally.

The mentioned shortcomings encourage the formulation of CC so that it is variational and also applicable to correlated systems. This gives rise to the Unitary Coupled Cluster (UCC) method. It consists of redefining the ansatz I.133 to

$$|\psi_{UCC}\rangle = e^{T-T^\dagger} |\psi_{HF}\rangle. \quad (\text{I.137})$$

Obviously, the operator $T - T^\dagger$ is anti-Hermitian, therefore e^{T-T^\dagger} is unitary and the formulation is therefore variational:

$$E = \min_t \langle \psi_{HF} | e^{-(T-T^\dagger)} H e^{T-T^\dagger} | \psi_{HF} \rangle. \quad (\text{I.138})$$

Unfortunately, an ansatz of the form I.137 does not lead to equations that would be solvable on a classical computer [61] (BCH expansion is not finite). Fortunately, the application of unitary operators is a natural domain of quantum computers.

The implementation of UCC on a quantum computer can be divided into two steps:

1. Preparation of the reference state $|\Phi_0\rangle \equiv |\psi_{HF}\rangle$, which in this work is always considered as the result of the Hartree-Fock method
2. Application of the UCC operator $U(\theta) \equiv U(t)$, where in this case the general set of parameters θ is represented by the set of amplitudes t from the expression I.132

The algorithm starts with an initial estimate of the amplitudes $t^{(0)}$, which then iteratively converge to a resulting set $t^{(n)}$ through energy minimization, providing a good approximation of the ground state. These steps will be discussed in more detail.

The reference state is required to have the largest possible overlap with the resulting state and to be easily implementable. In many cases, the H-F state (composed of a single determinant) is entirely sufficient. Its implementation is trivial. When working in the Jordan-Wigner basis, it's a state of the form $|0\rangle^{\otimes M-N} \otimes |1\rangle^{\otimes N}$, where M is the number of considered MOs and N is the number of electrons in the system. In the case of parity or B-K bases, the order of zeros and ones will be swapped, but it will still be a state of the computational basis. For strongly correlated systems, the H-F state is not a good initial estimate. A well-chosen initial state can then be obtained, for example, by the MCSCF method. It can be shown [58] that as long as such a state contains only a polynomial number of computational basis states, it is efficiently implementable.

Implementing $U(\mathbf{t})$ might seem more challenging. However, the truth is that most of the necessary theory has already been discussed in this work. The UCC operator can be written as

$$U(\mathbf{t}) = e^{\sum_j t_j (\tau_j - \tau_j^\dagger)}, \quad (\text{I.139})$$

where τ_j is the excitation operator and t_j the corresponding amplitude. The Trotter-Suzuki formula discussed in the Trotterization chapter can be applied to this expression

$$U(\mathbf{t}) \approx \left(\prod_j e^{\frac{t_j}{S} (\tau_j - \tau_j^\dagger)} \right)^S. \quad (\text{I.140})$$

Additionally, the advantage of using trotterization in VQE compared to its use in QPE is that the resulting errors can be effectively compensated within the classical optimization process [68]. In [51], it is shown by numerical calculation that the choice of ansatz $S = 1$ or $S = 2$ is as effective as if it were based directly on I.139. In the following, therefore, only $S = 1$ will be considered for simplicity, that is

$$U_1(\mathbf{t}) = \prod_j e^{t_j (\tau_j - \tau_j^\dagger)}. \quad (\text{I.141})$$

Mapping to qubits (either J-W or B-K) yields

$$(\tau_j - \tau_j^\dagger) = i \sum_k^{2^{l_j-1}} P_k^j, \quad (\text{I.142})$$

where P_i^j represents a Pauli chain with a real coefficient and the index k goes through 2^{l_j-1} of these chains, where l_j denotes the order of excitation of the τ_j operator. It can be shown [51] that all strings formed from one expression $(\tau_j - \tau_j^\dagger)$ commute with each other. The resulting form is

$$U_1(\mathbf{t}) = \prod_j \prod_k^{2^{l_j-1}} e^{it_j P_k^j}, \quad (\text{I.143})$$

which of course can be easily implemented on a quantum computer.

It is good to mention that although B-K mapping appears to be more advantageous than J-W, on architectures with limited connectivity (such as superconducting qubits) the necessity of subsequent application of SWAP operations may completely cancel this advantage. Of course, in order to achieve a polynomial number of operations, the ansatz (as with classic CC) must be shortened and consider e.g. UCC Single and Double (UCCSD), where

$$T \approx T_1 + T_2. \quad (\text{I.144})$$

When choosing UCCSD, the number of parameters increases as $\binom{M-N}{2}\binom{N}{2} + \binom{M-N}{1}\binom{N}{1} < O(M^2N^2)$.

The total number of gates then increases as $O(\log(M)M^2N^2)$ in case of B-K mapping and as $O(M^3N^2)$ in case of J-W mapping [51]. Such scaling in the number of parameters represents a great challenge for subsequent classical optimization, and it is therefore necessary to use various approximate techniques known from classical computational chemistry, such as limiting the calculation to the active space or freezing internal orbitals. Several optimization techniques have been tried in the context of UCC, each showing varying degrees of robustness to errors arising from device imperfections. They will be discussed in more detail in later sections.

I.8.2 Hardware Efficient Ansatz

Hardware Efficient Ansätze (HEA) is a specific representative of the class of hardware-heuristic ansätze. The origins of this approach can be found in [46], but the work [30] became the basis for subsequent experimental implementations (primarily on superconducting platforms). In preparing the heuristic trial state, two types of operations are utilized. The first are simple single-qubit Euler rotations $U(\theta)$, and the second are entangling blocks U_{ent} acting on pairs of qubits (thus, a fixed set of two-qubit gates). U_{ent} is chosen according to the specifics of the particular processor. Natural interactions on the given device can be described by a so-called *drift* Hamiltonian H_0 , from which then $U_{ent} = \exp\{-iH_0\tau\}$. The parameterized trial state is then obtained by applying D entangling blocks interspersed with Euler rotations to the initial (commonly H-F) state

$$|\psi_{HEA}(\theta)\rangle = \underbrace{U^{(D)}(\theta^{(D)})U_{ent}\dots U^{(1)}(\theta^{(1)})U_{ent}}_{D\text{-times}} U^{(0)}(\theta^{(0)}) |\psi_{HF}\rangle. \quad (\text{I.145})$$

One-qubit Euler rotations can be written (from the well-known Z-X decomposition) as

$$U^{(k)}(\theta^{(k)}) = \bigotimes_{i=0}^{M-1} \left(R_i^Z(\theta_{i,1}^{(k)}) R_i^X(\theta_{i,2}^{(k)}) R_i^Z(\theta_{i,3}^{(k)}) \right), \quad (\text{I.146})$$

where the index i goes over all qubits. Since the initial rotation around the z axis has no effect on the states of the computational basis, it can be omitted, and the total number of independent angles is thus $p = M(3D + 2)$.

It is good to mention that the parameter τ (evolution time) in the expression for U_{ent} is also controllable. However, as shown in [30], the best optimization results are achieved for a fixed τ . Since here (unlike in UCC) it does not depend on the implementation of specific entangling operations, but rather on U_{ent} generating sufficient entanglement, τ is chosen (based on numerical calculations) to minimize decoherence effects.

It can be noticed that the entire process mentioned so far does not inherently have the property of conserving the number of particles. In this basic scheme, therefore, this requirement must be considered during the classical optimization process. In light of this problem, entangling blocks that conserve the number of particles have been proposed [6] and are thus more suitable for simulating the electronic structure. Moreover, they are very naturally implementable on superconducting architectures.

In the mentioned work, the entangling blocks represent two types. The first, $U_{ent}^{(1)}$, consists of two-parameter particle-number conserving gates of the form

$$U_1(\theta_1, \theta_2) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta_1 & e^{i\theta_2} \sin \theta_1 & 0 \\ 0 & e^{-i\theta_2} \sin \theta_1 & -\cos \theta_1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{I.147})$$

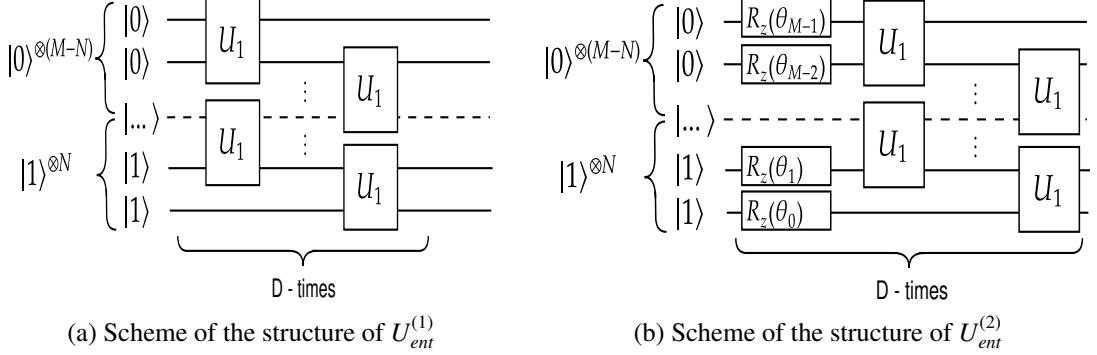


Figure I.5

The second option is to assemble the entangling block $U_{ent}^{(2)}$ from the one-parameter, number of particles preserving gates of the form

$$U_2(\theta) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos 2\theta & -i \sin 2\theta & 0 \\ 0 & -i \sin 2\theta & \cos 2\theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (\text{I.148})$$

which can be seen as a special case of U_1 . A schematic drawing of the circuit structure for these two types of blocks is shown in the figure I.5. The implementation of gates U_1 and U_2 using elementary gates is trivial and is detailed in [6]. The suitable choice of classical optimization technique will be discussed in subsequent sections.

As has been mentioned, purely hardware-heuristic ansätze, due to barren plateaus in the Hilbert space where classical optimization occurs, appear to be rather unscalable solutions. It is therefore important to mention that there are other ansätze with various degrees of physical and hardware motivation. Among the most important are the *Hamiltonian Variational Ansatz* and the *Low Depth Circuit Ansatz*. Another very interesting idea that connects the field of quantum simulations and quantum machine learning is to use so-called autoencoders for the compression of quantum data (states), i.e., for their representation using a smaller number of qubits. Here, the autoencoder serves to find a unitary operation that coherently maps quantum data onto this smaller number of qubits. This subsequently allows for the use of the so-called *Compressed Unsupervised State Preparation* (CUSP) method, which can generate a new ansatz that will be shallower in terms of the number of gates and thus less prone to errors [52].

I.8.3 Energy measurement

It is clear that measuring the energy of a given state could be done using QPE, but due to the previously mentioned problems with coherence times, it does not make sense to discuss this option within the context of VQE. Instead, another method of measuring energy was proposed, namely the so-called *Hamiltonian averaging* [38]. This is based on the fact that the Hamiltonian is written as a sum of Pauli strings

$$H = \sum_j h_j P_j = \sum_j h_j \bigotimes_i \sigma_i^j, \quad (\text{I.149})$$

where h_j are the coefficients resulting from the calculation of molecular integrals and σ_i^j denotes one of the Pauli matrices X, Y, Z or I , while i indexes the qubits on which they act. The basis is the use of the

linearity of the expected value

$$E(\boldsymbol{\theta}) = \langle H \rangle(\boldsymbol{\theta}) = \langle H \rangle_{|\psi(\boldsymbol{\theta})\rangle} = \sum_j h_j \langle \psi(\boldsymbol{\theta}) | \bigotimes_i \sigma_i^j | \psi(\boldsymbol{\theta}) \rangle. \quad (\text{I.150})$$

It is evident that each term in the expression I.149 is measured separately, and after each measurement, the parameterized state with the same parameters must be prepared again until measurements have been made for all terms. Only then are the parameters updated. However, it is necessary to view the measured Pauli strings as random variables and accordingly adjust the number of measurements for these individual strings [40], [38]. First, it's good to realize that measuring the expected value of a Pauli matrix X or Y corresponds to a change in the computational basis, specifically, it suffices to apply $R_x(-\frac{\pi}{2})$ or a Hadamard gate before measurement, respectively. Furthermore, the eigenvalues of Pauli matrices are ± 1 . This remains valid even for tensor products of Pauli matrices. Thus, measured values on one term of the Hamiltonian expansion $h_j P_j$ will be $\pm h_j$. The number of measurements that should be performed on a given term depends on the variance of the quantity (given string) on the prepared parameterized state. The variance of a quantity (and specifically of Pauli strings) is given as

$$\sigma^2(P_j) = \langle P_j^2 \rangle - \langle P_j \rangle^2 = 1 - \langle P_j \rangle^2, \quad (\text{I.151})$$

where in the last equality the fact that exponentiation of the tensor product of the Pauli matrices gives unity is used. If the j th term of the Hamiltonian is measured m_j times, then the statistical error ϵ_j associated with this term is given by [51]

$$\epsilon_j^2 = \frac{|h_j|^2 \sigma^2(P_j)}{m_j} \leq \frac{|h_j|^2}{m_j}. \quad (\text{I.152})$$

The goal is to measure the expected energy with an accuracy of ϵ . All strings are assumed to be measured independently (i.e., $\text{Cov}(\langle P_i \rangle, \langle P_j \rangle) = 0, \forall i \neq j$), and thus the total variance will then be determined by the sum of the variances of the individual members. From this it is clear that it is sufficient to choose $\epsilon_j^2 = \frac{|h_j|}{\sum_i |h_i|} \epsilon^2$. From this already follows the total required number of measurements m

$$m = \sum_j m_j = \sum_j \frac{|h_j| \sigma^2(P_j) \sum_i |h_i|}{\epsilon^2} \leq \frac{\left(\sum_j |h_j| \right)^2}{\epsilon^2} \quad (\text{I.153})$$

From this, it is evident that the number of measurements depends on the magnitudes of the coefficients in the Hamiltonian, which are determined by molecular integrals, further determined by the choice of basis set.

As previously derived, the number of terms in the Hamiltonian of the electronic structure grows with the number of orbitals M as $O(M^4)$. From the expression I.153, it might seem that $O(M^8/\epsilon^2)$ measurements would be necessary. However, utilizing the locality of basis sets, it was shown for GTO [38] that the number of terms in the Hamiltonian whose overlap integrals give non-negligible contributions grows only as $O(M^2)$, and advanced techniques can evaluate these contributions in time $O(M)$. This analysis applied to Hamiltonian averaging showed that to achieve accuracy ϵ , $O(M^6/\epsilon^2)$ measurements are needed. Estimates of the number of measurements were also made for the case of plane wave bases [5].

The overall computational effort is determined by the Hamiltonian averaging, the complexity of the state preparation ansatz, and the necessary number of evaluations of the parameterized expected value required by the optimization process. The following paragraphs will be dedicated to methods for reducing the number of measurements in Hamiltonian averaging.

The first method is called *term truncation* [40]. In short, the idea of this method is to omit (i.e., not measure at all) terms whose values contribute minimally to the overall expected value. This can be achieved by sorting the Pauli strings within the Hamiltonian according to their maximum contribution to the expected energy. This is given as $|\langle h_j P_j \rangle| \leq |h_j|$. Thus, terms can be sorted from smallest to largest by $|h_j|$. Subsequently, a sequence of partial sums

$$s_k = \sum_{i=1}^k |h_i| \quad s_0 = 0, \quad (\text{I.154})$$

which determine the maximum distortion caused by truncating the first k terms. Then, it is sufficient to choose a constant $C \in [0, 1)$ and remove the first k' terms such that in the given sequence, the maximum index k' is found for which $s_{k'} \leq C\epsilon$. Since C represents the distortion of the expected value, it must be reflected in the number of measurements to achieve accuracy ϵ . To preserve the accuracy ϵ in the result, the variance on the remaining terms must be reduced so that it will satisfy

$$C^2\epsilon^2 + \sum_i^{O(M^4)-k'} \sigma^2(\langle P_i \rangle) < \epsilon. \quad (\text{I.155})$$

It can be seen that this is satisfied by choosing $\sigma^2(\langle P_i \rangle) = (1 - C^2)\epsilon^2/(M - k')$ (in other words, that the remaining terms need to be calculated with an accuracy of $(1 - C^2)\epsilon^2$). The expected number of measurements will then be given as

$$m_{tr} = \frac{\sum_i^{O(M^4)-k'} |h_i| \sigma^2(\langle P_i \rangle)}{(1 - C^2)\epsilon^2}. \quad (\text{I.156})$$

The constant C is chosen to maximize computational efficiency for a specific experiment.

Another method is the grouping of commuting terms in the Hamiltonian [40] (which can be done using a classical sorting algorithm with complexity $O(M^2)$). If P_i and P_j commute, then they can be measured sequentially on the same state without distorting the resulting expected value. Such measurement can be performed, for example, with the help of an ancilla qubit and adding a small number of gates. Since generally, state preparation is much more demanding than projective measurement, this method offers significant savings in the number of gates and the errors that stem from them. However, since the terms within a commuting set are measured on the same state, their covariance (pairwise) may no longer be zero. This, in analogy with simulations using the classical Monte Carlo method [24], requires a larger or smaller number of measurements depending on whether the covariance is positive or negative. This is why it is necessary to carefully choose into which commuting sets to divide the operators. An example of the impact of division into different sets for a simple Hamiltonian is given in [40]. Since the covariance between terms is dependent on the state on which it is evaluated, classical algorithms sorting terms into commuting sets will benefit from the best classical approximation of the ground state.

Another very interesting method for reducing the number of measurements is the use of conditions of so-called *N-representability* placed on the reduced density matrices of the system [54]. Detailed derivation and description of these conditions are quite complex and beyond the scope of this work. In short, however, these conditions provide additional information about the system, which can be used to reduce the number of measurements. These conditions are incorporated into the calculation in the form of terms added to the Hamiltonian, with the resulting Hamiltonian having the same expected value as the original.

Finally, it is worth mentioning the Bayesian approach to estimating expected values of the Hamiltonian briefly discussed in [40].

I.8.4 Classical optimization

In classical optimization methods, it is necessary to consider speed, accuracy, and for NISQ devices, their resilience to noise as well. Meeting these three criteria can be particularly challenging for optimization in a multidimensional parametric space, and different methods may be more or less advantageous for different problems.

Broadly speaking, optimization algorithms can be divided into *direct search* and those based on *gradient descent*. While direct search usually proves to be more resistant to noise, on the other hand, it may require a larger number of evaluations of the cost function [32]. Algorithms based on the gradient of the cost function can further be divided according to whether they use only the first or also the second derivatives, into first-order and second-order optimizers. This work will briefly describe a representative of each of these three categories.

I.8.4.1 Direct search - Nelder-Mead algorithm

Direct search refers to optimization methods that do not use derivatives of the cost function. A selected representative of this broad class for this work is the Nelder-Mead algorithm [41]. This method is based on a simplex, which can intuitively be described as a generalization of a triangle to any dimension. Mathematically, a simplex \mathcal{S} in \mathbb{R}^k is defined as the convex hull of $k + 1$ affinely independent points.

The idea is that a random simplex is first generated, and then the values of the cost function at each of the vertices are iteratively reduced. Specific implementations of the Nelder-Mead algorithm may vary slightly; here, the most important steps will be discussed [62].

1. Sort the values of the cost function at individual points by size:
 $E(\theta_1) \leq E(\theta_2) \leq \dots \leq E(\theta_{k+1})$ and check if the conditions for terminating the algorithm are met.
2. Compute centroid θ_0 of all points except θ_{k+1}
3. Calculate the reflection point $\theta_r = \theta_0 + \alpha_r(\theta_0 - \theta_{k+1})$ where $\alpha_r > 0$ is the reflection coefficient.
4. If:
 - (a) θ_r is not the best guess among all points, but it is better than the second worst point θ_k , then replace θ_{k+1} by θ_r and go back to step 1.
 - (b) θ_r is the best of all points, then calculate the expansion point $\theta_e = \theta_0 + \alpha_e(\theta_r - \theta_0)$, where $\alpha_e > 1$ is the expansion coefficient, replace the worst point θ_{k+1} with the better of the points θ_r, θ_e and return to step 1.
5. At this point $E(\theta_r) \geq E(\theta_k)$. If:
 - (a) $E(\theta_r) < E(\theta_{k+1})$, calculate the outer contraction point $\theta_c = \theta_0 + \alpha_c(\theta_r - \theta_0)$, where $0 < \alpha_c < 0.5$ is the contraction coefficient and if θ_c better than θ_r , so replace θ_{k+1} with θ_c and go to step 1.
 - (b) $E(\theta_r) \geq E(\theta_{k+1})$, calculate the internal contraction point $\theta_c = \theta_0 + \alpha_c(\theta_{k+1} - \theta_0)$, where $0 < \alpha_c < 0.5$ is the contraction coefficient and if θ_c better than θ_{k+1} , so replace θ_{k+1} with θ_c and go to step 1 .
6. (Occurs in the rare case that contraction increases the value of the cost function) Replace all points except the best one (θ_1) with points $\theta_i = \theta_1 + \alpha_s(\theta_i - \theta_1)$ where α_s is the shrinkage coefficient.

It is evident that this algorithm requires the evaluation of only a few points at each iteration, making it very friendly for VQE from this perspective. However, it is necessary to mention that even within the class of direct search algorithms, there are algorithms that appear to be more efficient. Examples can include algorithms from TOMLAB [40].

I.8.4.2 Gradient-based methods

Gradient calculation

A large number of optimization methods are based not on evaluating the cost function at given points but on evaluating the gradient of this function. There are two main ways to calculate the gradient of a function. The first are stochastic approximation methods, and the second is the analytical calculation of the gradient.

A representative of the first class is the so-called *finite difference stochastic approximation (FDSA)*. Here, the approximation of the gradient at point $\theta^{(k)}$ is given simply as

$$(\nabla E(\theta^{(k)}))_j = \frac{E(\theta^{(k)} + c_k e_j) - E(\theta^{(k)} - c_k e_j)}{2c_k}. \quad (\text{I.157})$$

On the other hand, in the analytical calculation of the gradient, the measurement of the observable gradient is performed directly on the quantum computer using the analytical properties of the ansatz. It is good to recall that the variational ansatz $U^{(k)}(\theta^{(k)})$ generally has the form

$$U^{(k)}(\theta^{(k)}) = \prod_j U_j^{(k)}(\theta_j^{(k)}) = \prod_j \exp\{i\theta_j^{(k)} P_j^{(k)}\}, \quad (\text{I.158})$$

where $P_j^{(k)}$ is a Hermitian matrix in general, commonly a Pauli chain. It is then not difficult to deduce that the partial derivative of the expected value of some observable $O^{(k)}$ (in our case usually Pauli chains as terms of the Hamiltonian) with respect to $\theta_j^{(k)}$ will be

$$\frac{\partial \langle O^{(k)}(\theta^{(k)}) \rangle}{\partial \theta_j} = 2 \operatorname{Im}\left\{ \langle \psi_0 | V_j^{(k)\dagger}(\theta^{(k)}) M^{(k)} U^{(k)}(\theta^{(k)}) | \psi_0 \rangle \right\}, \quad (\text{I.159})$$

where $M^{(k)}$ is the measurement operator (commonly the Pauli chain according to the measured term of the Hamiltonian) and $V_j^{(k)}(\theta^{(k)})$ is essentially the derivative of the operator $U_j^{(k)}(\theta^{(k)})$, i.e.

$$V_j^{(k)}(\theta^{(k)}) = e^{i\theta_p^{(k)} P_p^{(k)}} \dots e^{i\theta_j^{(k)} P_j^{(k)}} e^{i\theta_{j-1}^{(k)} P_{j-1}^{(k)}} \dots e^{i\theta_1^{(k)} P_1^{(k)}}. \quad (\text{I.160})$$

The problem is that $V_j^{(k)\dagger}(\theta^{(k)}) M^{(k)} U^{(k)}(\theta^{(k)})$ is usually not Hermitian and thus there is no known method to convert it to a quantum circuit. However, this can be easily circumvented by e.g. adding an ancilla qubit [62].

First order optimizers

The basic first-order method is the *simple* (sometimes also referred to as *vanilla*) *gradient descent*. It is based on the fact that the negatively taken gradient of the function at a given point indicates the direction of its steepest descent. The simplest scheme of this algorithm can thus be summarized as:

1. Check whether the conditions for the termination of the algorithm (convergence) are not met.
2. Calculate the gradient $\nabla E(\theta^{(k)})$

3. Generate a new set of parameters as $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} - \eta \nabla E(\boldsymbol{\theta}^{(k)})$, where η denotes the size of the optimization step - the so-called *learning rate*, and go to step 1.

To this basic scheme, of course, there are many enhancements, such as optimizing η between steps and others. The gradient here can be calculated using the previously mentioned approximate or analytical methods. In either case, however, it is evident that calculating the entire gradient requires potentially a large number of partial derivatives, which is not too favorable from the VQE perspective.

For this reason, another first-order method, *Simultaneous Perturbation Stochastic Approximation (SPSA)* [59], is much more commonly used within VQE. It involves a stochastic approximation of the gradient, where instead of calculating all the partial derivatives, a random perturbation vector is generated at each iteration, along which the gradient is approximated - descending according to a randomly chosen direction. The formula for this approximation is

$$(\nabla E(\boldsymbol{\theta}^{(k)}))_j = \frac{E(\boldsymbol{\theta}^{(k)} + c_k \Delta^{(k)}) - E(\boldsymbol{\theta}^{(k)} - c_k \Delta^{(k)})}{2c_k (\Delta^{(k)})_j}, \quad (\text{I.161})$$

where $\Delta^{(k)}$ is the perturbation vector. It is evident that only two evaluations of the cost function are necessary at each iteration, which is very favorable for VQE. Unlike simple gradient descent, SPSA has been successfully used to find the ground state due to its lower susceptibility to noise [23], [30].

Second order optimizers

Second-order methods use the second derivatives of the cost function to determine the direction of descent. Perhaps the most utilized second-order algorithm within VQE is the *Broyden-Fletcher-Goldfarb-Shanno (BFGS)* algorithm [13] or its derived versions such as *Limited-memory BFGS Bound (L-BFGS-B)* [14]. The direction of descent is determined by preconditioning the gradient with curvature information, i.e., its modification using the Hessian matrix. Given that computing the full Hessian matrix (used in the original BFGS algorithm) is very demanding, L-BFGS uses an approximation of the Hessian matrix. This can be calculated, for example, using SPSA [62]. The basic BFGS algorithm can be summarized in steps:

1. Compute the approximate Hessian matrix $B^{(0)}$ from the initial guess $x^{(0)}$
2. Repeat until the conditions for termination (convergence) are met:
 - (a) Obtain the descent direction of $\mathbf{p}^{(k)}$ by solving the equation $B^{(k)} \mathbf{p}^{(k)} = -\nabla E(\boldsymbol{\theta}^{(k)})$
 - (b) perform a *line search* (minimization on the line) to find $\alpha^{(k)} = \text{argmin}(E(\boldsymbol{\theta}^{(k)} + \alpha \mathbf{p}^{(k)}))$, where $\alpha \in \mathbb{R}^+$.
 - (c) Set $\mathbf{s}^{(k)} = \alpha^{(k)} \mathbf{p}^{(k)}$ and update $\boldsymbol{\theta}^{(k+1)} = \boldsymbol{\theta}^{(k)} + \mathbf{s}^{(k)}$
 - (d) Set $\mathbf{y}^{(k)} = \nabla E(\boldsymbol{\theta}^{(k+1)}) - \nabla E(\boldsymbol{\theta}^{(k)})$
 - (e) Set $B^{(k+1)} = B^{(k)} + \frac{\mathbf{y}^{(k)} (\mathbf{y}^{(k)})^T}{(\mathbf{y}^{(k)})^T \mathbf{s}^{(k)}} - \frac{B^{(k)} \mathbf{s}^{(k)} (\mathbf{s}^{(k)})^T (B^{(k)})^T}{(\mathbf{s}^{(k)})^T B^{(k)} \mathbf{s}^{(k)}}.$

Chapter II

Practical part

In this section, a specific example of a molecule - molecular hydrogen H₂ - will be analyzed in detail. Due to the small size of the molecule, a large part of the derivation can be done quite illustratively without the use of computational technology. First, mapping to qubits for different choices of basis sets will be shown. On this occasion, it will also be demonstrated how the necessary number of qubits can be reduced and the so-called *active space* limited by using symmetries. Then, the actual implementation in the open-source framework for quantum computing Qiskit [2] will follow. Given the limitations of NISQ computers, only the approach using VQE will be considered in the implementation. First, a high-level view of sampling the potential energy surface (PES) with different choices of basis sets and mappings will be discussed. Then, the results of VQE with different choices of basis sets, ansatze, and classical optimizers will be compared in more detail. The implementation will take place in a Jupyter Notebook environment and will utilize, in addition to Qiskit documentation, materials and codes from the IBM Quantum Spring Challenge 2022 and Qiskit Global Summer School 2022, which the author participated in and successfully completed.

II.1 H₂ - mapping and reduction of the number of qubits

In deriving the mapping from the very beginning, primarily the materials from [55] and [37] will be utilized. The discussion will then be divided according to the used basis set. Special attention will be given to the STO-3G set for its illustrative purposes.

By specifying the electronic Hamiltonian for molecular hydrogen, which has two electrons and two nuclei, it turns out

$$H_{H_2} = - \sum_{i=1}^2 \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{i,I=1}^2 \frac{e^2}{4\pi\varepsilon_0} \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \frac{1}{2} \sum_{i,j=1, i \neq j}^2 \frac{e^2}{4\pi\varepsilon_0} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}. \quad (\text{II.1})$$

To express this Hamiltonian in second quantization using relations I.26, I.27, I.28, a basis set must already be chosen in which the molecular integrals will be calculated. It must be kept in mind that the more comprehensive the basis set, the better it will approximate the real MOs, and the lower the ground state energy values that can be achieved.

II.1.1 STO-3G

The STO-3G basis for molecular hydrogen (also called the minimal basis) contains only the {1s} orbital for each atom, represented by a linear combination of three Gaussian orbitals. Considering

two possible spins in each orbital, a total of 4 qubits will be needed. Let's denote these orbitals as $|1s_{A\uparrow}\rangle, |1s_{A\downarrow}\rangle, |1s_{B\uparrow}\rangle, |1s_{B\downarrow}\rangle$. Here, A, B denote the atoms around whose nuclei the orbital is centered, and the arrows indicate the value of s_z spin projection in the orbital. It can be shown that for molecular hydrogen, the basis of molecular orbitals (MOs) is given as

$$|\sigma_{g\uparrow}\rangle = \frac{1}{\sqrt{N_g}}(|1s_{A\uparrow}\rangle + |1s_{B\uparrow}\rangle) \quad |\sigma_{g\downarrow}\rangle = \frac{1}{\sqrt{N_g}}(|1s_{A\downarrow}\rangle + |1s_{B\downarrow}\rangle) \quad (\text{II.2})$$

$$|\sigma_{u\uparrow}\rangle = \frac{1}{\sqrt{N_u}}(|1s_{A\uparrow}\rangle - |1s_{B\uparrow}\rangle) \quad |\sigma_{u\downarrow}\rangle = \frac{1}{\sqrt{N_u}}(|1s_{A\downarrow}\rangle - |1s_{B\downarrow}\rangle), \quad (\text{II.3})$$

where $N_g = 2(1 + \langle 1s_A | 1s_B \rangle)$ and $N_u = 2(1 - \langle 1s_A | 1s_B \rangle)$ are the overlap-dependent normalization factors. Slater's determinant in the occupation numbers basis will be

$$|\psi\rangle = |n_{\sigma_{u\downarrow}}, n_{\sigma_{u\uparrow}}, n_{\sigma_{g\downarrow}}, n_{\sigma_{g\uparrow}}\rangle. \quad (\text{II.4})$$

Here, $n_k = 0$ if the orbital is empty and $n_k = 1$ if it is occupied. This representation precisely corresponds to the state in Jordan-Wigner (J-W) mapping.

Using these orbitals, molecular integrals are calculated, which is a standard procedure. Specifically, in this work, the computational chemistry package Psi4 [65] will be used due to its compatibility with the Windows system. Next, it is necessary to map fermionic creation and annihilation operators onto strings of Pauli matrices. While JW and parity mappings are straightforward and can be done on paper for the molecule of hydrogen in the minimal basis, Bravyi-Kitaev (B-K) mapping requires more attention and even for the specific case of H₂ is detailed in [55]. These processes are, of course, already automated, and Qiskit Nature module will be used for further work. Qubit Hamiltonians in individual mappings can be written as:

$$\begin{aligned} H_{H_2,JW} = & h_0I + h_1Z_0 + h_2Z_1 + h_3Z_2 + h_4Z_3 + h_5Z_0Z_1 + h_6Z_0Z_2 + h_7Z_1Z_2 + h_8Z_0Z_3 \\ & + h_9Z_1Z_3 + h_{10}Z_2Z_3 + h_{11}Y_0Y_1X_2X_3 + h_{12}X_0Y_1Y_2X_3 + h_{13}Y_0X_1X_2Y_3 + h_{14}X_0X_1Y_2Y_3, \end{aligned} \quad (\text{II.5})$$

$$\begin{aligned} H_{H_2,Par} = & h_0I + h_1Z_0 + h_2Z_0Z_1 + h_3Z_1Z_2 + h_4Z_3Z_2 + h_5Z_1 + h_6Z_0Z_1Z_2 + h_7X_0X_2Z_3 + h_8X_0Z_1X_2 \\ & + h_9X_0Z_1X_2Z_3 + h_{10}X_0X_2 + h_{11}Z_0Z_2Z_3 + h_{12}Z_0Z_2 + h_{13}Z_0Z_1Z_2Z_3 + h_{14}Z_1Z_3, \end{aligned} \quad (\text{II.6})$$

$$\begin{aligned} H_{H_2,BK} = & h_0I + h_1Z_0 + h_2Z_0Z_1 + h_3Z_2 + h_4Z_1Z_2Z_3 + h_5Z_1 + h_6Z_0Z_2 + h_7X_0X_2Z_3 + h_8X_0Z_1X_2 \\ & + h_9X_0Z_1X_2Z_3 + h_{10}X_0X_2 + h_{11}Z_0Z_1Z_2Z_3 + h_{12}Z_0Z_1Z_2 + h_{13}Z_0Z_2Z_3 + h_{14}Z_1Z_3. \end{aligned} \quad (\text{II.7})$$

The H-F state in J-W mapping is then simply

$$|\psi_{H_2,JW}^{HF}\rangle = |0011\rangle, \quad (\text{II.8})$$

which represents the Slater determinant

$$\psi_{H_2}^{HF}(\mathbf{r}_1, \mathbf{r}_2) = \frac{1}{\sqrt{2}}(\sigma_{g\uparrow}(\mathbf{r}_1)\sigma_{g\downarrow}(\mathbf{r}_2) - \sigma_{g\uparrow}(\mathbf{r}_2)\sigma_{g\downarrow}(\mathbf{r}_1)). \quad (\text{II.9})$$

It can be shown by more precise classical methods [27] that the state better corresponding to the true ground state (while keeping the same value of spin projection and number of electrons) is

$$|\psi_{H_2,JW}^g\rangle = 0,9939|0011\rangle - 0,1106|1100\rangle. \quad (\text{II.10})$$

This shows that the H-F state is a very good approximation of the ground state. The second determinant appears due to dynamical correlations between electrons due to their electrostatic repulsion.

II.1.2 6-31G

The 6-31G basis set has a double-zeta representation of valence electrons, i.e., for each atom, we must consider 4 spin orbitals $\{1s_{\uparrow}, 1s_{\downarrow}, 1s'_{\uparrow}, 1s'_{\downarrow}\}$. Thus, a total of 8 qubits will be needed. This example vividly demonstrates how, in B-K mapping, the conservation of particle number and total spin projection can be used to reduce the number of qubits. The basis considered will again be the MOs obtained using the H-F method.

Let the arrangement of MOs be such that the first four have a positive spin projection and the remaining four have a negative spin projection. In J-W mapping, the H-F state will thus be given as $|00010001\rangle$. Using the recursive relation I.101, it is evident that the transformation matrix for transitioning from J-W mapping to B-K mapping will have the form

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix}. \quad (\text{II.11})$$

Now, it is important to realize what the fourth and eighth rows of this matrix represent. Given the arrangement of MOs, the fourth row is the sum (mod 2) of all MO occupations with a positive spin projection, which equals the sum of electrons with a positive spin projection. Since this quantity is conserved, we can eliminate the fourth qubit. Similarly, it can be understood that the eighth row represents the sum of all particles, which is also conserved, and thus the eighth qubit can also be eliminated. When applying the full transformation matrix to the given state in J-W mapping, the state in B-K mapping is given as $|00111011\rangle$. After removing those two qubits, what remains is $|011011\rangle$.

II.1.3 cc-PVDZ

cc-PVDZ for H_2 includes a double-zeta representation for the valence shell plus additional higher polarization orbitals, which results in a set of $\{1s, 1s', 2p_x, 2p_y, 2p_z\}$ for each atom. For two atoms and two possible spins, a total of 20 qubits will be needed. This is quite a lot, and thus this basis set serves well to demonstrate computation in a so-called *reduced active space*.

It often happens that some orbitals are with very high probability occupied or, conversely, virtual in all Slater determinants that compose the wave function. Since the entire calculation of electronic structure is essentially an optimal distribution of electrons into orbitals, those orbitals whose occupancy is very likely not to change during the computation can be omitted from the calculation. The occupancy of an orbital can be determined using the formalism of reduced density matrices (RDMs) of the system [27]. The expected value of any one- or two-electron Hermitian operator O on the state $|\psi\rangle = \sum_n \alpha_n |n\rangle$ can be written as

$$\langle\psi|O|\psi\rangle = \sum_{i,j} O_{ij}\rho_{ij}^1 + \sum_{i,j,k,l} V_{ijkl}\rho_{ijkl}^2, \quad (\text{II.12})$$

where

$$\rho_{ij}^1 = \langle\psi|a_i^\dagger a_j|\psi\rangle, \quad \rho_{ijkl}^2 = \langle\psi|a_i^\dagger a_k^\dagger a_j a_l|\psi\rangle \quad (\text{II.13})$$

are elements of the so-called one-particle (1-RDM) and two-particle (2-RDM) density matrix, respectively. In this sense, the reduction using RDMs is considered with respect to the state that should be a

good approximation of the ground state and which can be obtained by a classical method. In the case of molecular hydrogen, for example, the Configurational Interaction Single and Double (CISD) method can be discussed. Importantly, the 1-RDM ρ^1 , whose diagonal elements are clearly the expected values of the particle number operator for a given orbital, will be primarily considered. ρ^1 is Hermitian and thus diagonalizable, which (when applying the same unitary matrix to orbitals as for diagonalization) represents a transition from the basis of canonical orbitals to so-called *natural molecular orbitals* (NMOs), and the diagonal elements in this basis are called *natural orbital occupation numbers* (NOONs). Orbitals with NOONs close to 0 or conversely 1 can thus be excluded.

It can be shown [37] that the diagonalized 1-RDM for molecular hydrogen, when using CISD, will have the form $\text{diag}(1, 96588 0, 00611 0, 02104 0, 0002 0, 00001 0, 00314 0, 00314 0, 00016 0, 00016 0, 00016)$, where both spin projections have been combined into one spatial orbital (otherwise, there would be 20 diagonal elements). Clearly, the fifth spatial orbital has by far the lowest occupancy probability and can be excluded from the calculation (considered always empty), which reduces the number of qubits from 20 to 18 by eliminating orbitals numbered as 10, 11. At this point, we consider alternating spin-up spin-down ordering.

Further reduction will again be possible through transition to B-K mapping using the BK-tree method, which precisely reflects the tree structure shown in figure I.1. First, we rearrange the orbitals so that all spin-up are first followed by all spin-down (0, 2, 4, 6, 8, 12, 14, 16, 18, 1, 3, 5, 7, 9, 13, 15, 17, 19). In exactly this order, we renumber the orbitals from 0 to 17. Next, the B-K transformation is performed exactly as described by this binary (Fenwick) tree with relationships for qubits described in I.97, just extended to 18 qubits. Now, just as for the 6-31G basis set, it can be noticed that the ninth qubit encodes the sum of all electrons with positive spin projection (mod 2), and the eighteenth qubit encodes the total number of electrons (mod 2), which are again conserved quantities, and thus these qubits can be excluded, resulting in a final count of 16 qubits.

II.2 Explicit UCCSD ansatz for STO-3G

This section will explicitly summarize how to construct the Unitary Coupled Cluster Ansatz with Single and Double Excitations (UCCSD) for molecular hydrogen. UCC is summarized by the relations I.137, I.131 and I.132. At the same time, UCCSD is limited to T_1 and T_2 only, so the operator for this ansatz has the form

$$U = e^{(T_1 - T_1^\dagger) + (T_2 - T_2^\dagger)}, \quad (\text{II.14})$$

where as a reminder

$$T_1 = \sum_{i \in ex, \alpha \in occ} t_{i\alpha} a_i^\dagger a_\alpha \quad T_2 = \sum_{i,j \in ex, \alpha, \beta \in occ} t_{ij\alpha\beta} a_i^\dagger a_j^\dagger a_\alpha a_\beta. \quad (\text{II.15})$$

Here occ denotes the occupied orbitals in the Hartree-Fock (H-F) state, and t are the parameters that are optimized. We consider the H-F state here to be $|0011\rangle$. Among all possible terms occurring in the sums for T_1 and T_2 , only those operators that preserve the total spin projection value of the molecule when applied to the H-F state are relevant. Specifically, these are only $a_2^\dagger a_0$, $a_3^\dagger a_1$, and $a_3^\dagger a_2^\dagger a_1 a_0$, and others relevant are equivalent to these. Thus, the UCCSD operator will take the form

$$U = e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2) + t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3) + t_{0123}(a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3)}. \quad (\text{II.16})$$

If one-step Trotterization is used, one can write

$$U \approx e^{t_{02}(a_2^\dagger a_0 - a_0^\dagger a_2)} e^{t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3)} e^{t_{0123}(a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3)}. \quad (\text{II.17})$$

It is not difficult to show that in the J-W mapping it holds

$$\begin{aligned} (a_2^\dagger a_0 - a_0^\dagger a_2) &= \frac{i}{2}(X_2 Z_1 Y_0 - Y_2 Z_1 X_0) \\ (a_3^\dagger a_1 - a_1^\dagger a_3) &= \frac{i}{2}(X_3 Z_2 Y_1 - Y_3 Z_2 X_1) \\ (a_3^\dagger a_2^\dagger a_1 a_0 - a_0^\dagger a_1^\dagger a_2 a_3) &= (X_3 Z_2 X_1 X_0 + Y_3 X_2 X_1 X_0 + Y_3 Y_2 Y_1 X_0 + Y_3 Y_2 X_1 Y_0 \\ &\quad - X_3 X_2 Y_1 X_0 - X_3 X_2 X_1 Y_0 - Y_3 X_2 Y_1 Y_0 - X_3 Y_2 Y_1 Y_0) \end{aligned}$$

As noted in [51], Pauli chains originating from the same excitation operators commute with each other, so one can write e.g.

$$e^{t_{13}(a_3^\dagger a_1 - a_1^\dagger a_3)} = e^{\frac{i t_{13}}{2} X_3 Z_2 Y_1} e^{-\frac{i t_{13}}{2} Y_3 Z_2 X_1}. \quad (\text{II.18})$$

It is interesting to mention that considerations valid only for a specific hydrogen molecule (generally non-scalable) can lead to the [28] UCCSD operator of the form

$$U = e^{-i\theta X_3 X_2 X_1 Y_0}. \quad (\text{II.19})$$

Implementing this operator is straightforward - for qubit 0, a basis rotation to y is used (using the gate $R_x(\frac{\pi}{2})$), and for the remaining three qubits, a basis rotation to x is used (using the Hadamard gate). Then, it suffices to sequentially entangle qubits 0 to 3 using CNOT gates and apply a rotation $R_z(2\theta)$ to qubit 3. Finally, the CNOT operations and basis rotations are reversed. By applying it to the H-F state, we obtain

$$U |0011\rangle = (\cos(\theta)I - i \sin(\theta)X_3 X_2 X_1 Y_0) |0011\rangle = \cos(\theta) |0011\rangle - \sin(\theta) |1100\rangle, \quad (\text{II.20})$$

from which the II.10 state can be obtained.

II.3 PES sampling under various choices of basis sets and mappings

In this section, the results of sampling the potential curve for H_2 will be discussed without a detailed examination of the internal setup of VQE. This is facilitated by the still very new and continuously developed Qiskit Nature module, which is designed so that, for example, chemists do not need to delve into the entire essence of quantum computing, but rather a few commands common to this field are sufficient, and everything else is resolved in the background. On the other hand, a more detailed view and analysis of VQE itself will be conducted in the following chapter.

Gradually, graphs comparing classical calculation and quantum computation simulation for the selected basis set across a wide range of nuclear distances will be presented. This calculation will always be performed in the parity mapping with reduction to 2 qubits due to computational speed. This graph will be followed by four graphs focusing on the dissociation curve in a much narrower range of nuclear distances in the area of the minimum at different types of mappings: J-W, parity without qubit reduction, parity with qubit reduction, and B-K. Only for the cc-PVDZ basis set, detailed graphs in the minimum area will not be presented, as very drastic limitations on the active space were necessary to make the calculation manageable on a regular notebook for a wide range of nuclear distances. The number of cost function evaluations will also be mentioned for each graph. Nuclear distances are given in Ångstroms and energy units in Hartree.

First, results for the STO-3G basis set will be presented. The calculation of the entire curve can be found in the attached material H2-PES-STO3G-PARITY (REDUCED) -WHOLE.ipynb and calculations of details in the minimum area in H2-PES-STO3G-ALL_MAPPINGS-DETAIL.ipynb

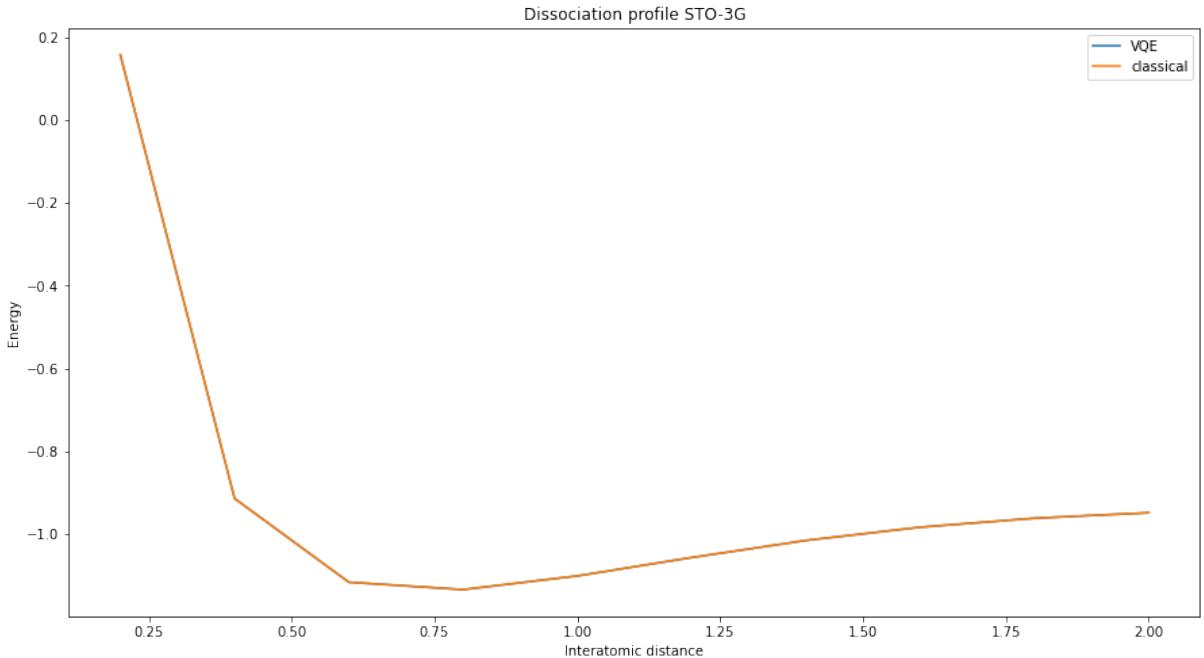


Figure II.1: The dissociation curve for molecular hydrogen using the STO-3G basis set and parity mapping with the reduction of two qubits. Samples calculated at ten equidistant points ranging from 0.2 to 2 Ångströms. The number of cost function evaluations: 366. Energy is given in units of Hartree, and at the lowest point, the VQE value is -1.1341475893375006 Hartree.

On the graph of the overall dissociation curve II.1, it can be seen that the results of the classical calculation match the VQE simulation so well that the classical calculation curve completely overlaps the VQE curve. In this sense, the accuracy of VQE is very good, and minor deviations will only be visible upon closer examination of the details. However, it must be kept in mind that the term accuracy here truly refers only to the difference compared to the classical calculation in the given basis set. STO-3G is the smallest possible set requiring the least number of qubits, hence the errors of VQE compared to the classical calculation are the smallest. However, the STO-3G set itself does not allow even the classical calculation to achieve as high precision (lower energy) as more comprehensive sets, which are computationally more demanding and require more qubits for simulation.

Detail in the area of the minimum for various mappings is shown in figure II.2. It is evident that the best results, both in terms of accuracy and in terms of the number of evaluations of the cost function, are achieved with parity mapping with reduction, where a value of -1.137305886339342 Hartree is reached at the lowest point. This demonstrates the advantage of reducing the search space and decreasing error rates by reducing the number of qubits. The worst accuracy results are achieved with B-K mapping, which indicates that the reduced non-locality of operators may not play a particularly crucial role for such small molecules.

Although the previous STO-3G example well demonstrates that in small search spaces, almost absolute agreement with the classical solution can be achieved, more comprehensive sets will be more interesting for comparison. For the 6-31G set case, the dissociation curve with parity mapping and the reduction of two qubits is illustrated in figure II.3, which was generated in the file H2-PES-631G-PARITY (REDUCED) -WHOLE.ipynb.

Here, a small deviation from the classical solution is already noticeable in the area of the minimum.

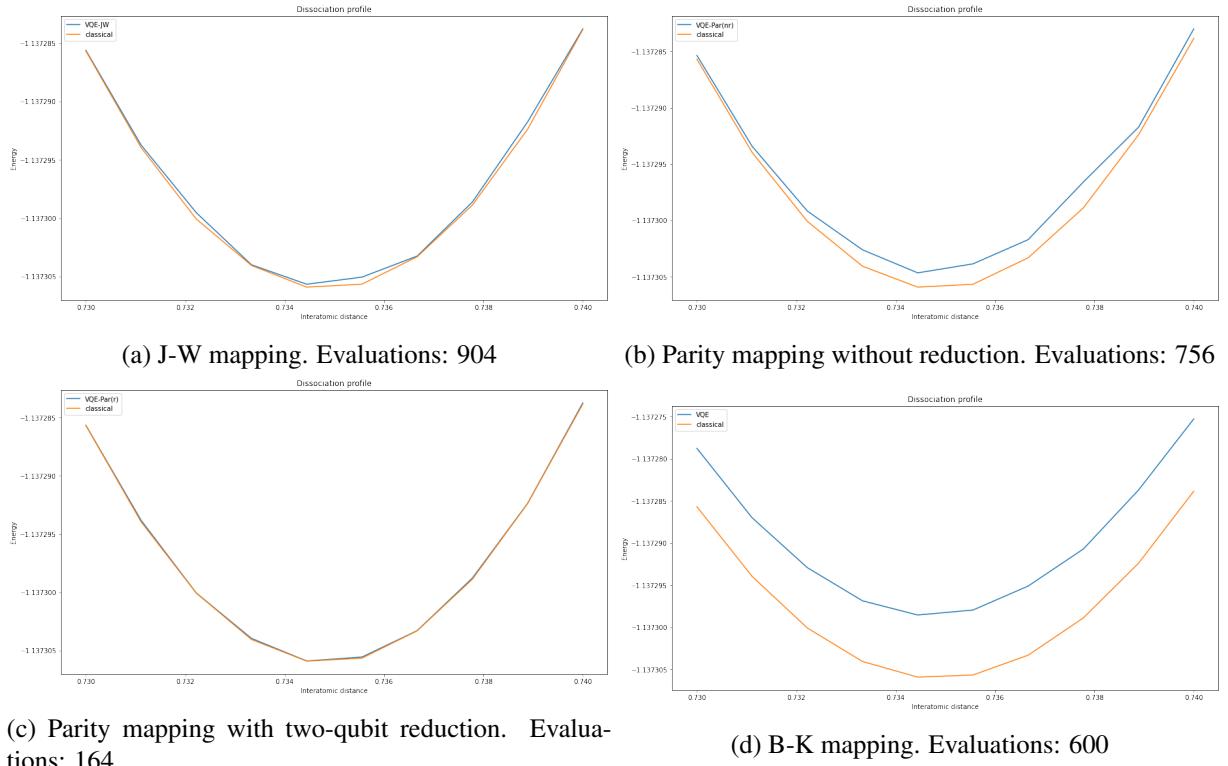


Figure II.2: Detail of the dissociation curves in the very narrow range of 0.73 to 0.74 Ångström for the STO-3G basis set in different mappings. Samples counted at ten equidistant points.

However, what is more important is the fact that even in this coarse sampling, the lowest energy is lower than what was observed in the detailed sampling for STO-3G. This demonstrates how a more comprehensive set is capable of better approximating the ground state. Detailed calculations in the area of the minimum for various mappings were conducted in the file `H2-PES-631G-ALL_MAPPINGS-DETAIL.ipynb`. The results are the graphs shown in figure II.4. From these graphs, it can be inferred that only J-W mapping and parity mapping with reduction provide relevant results in terms of accuracy. For B-K and parity mapping without reduction, not only is there a significant difference in energy determination, but the location of the minimum does not agree either. It is conceivable that one of the causes could be the optimizer getting caught on barren plateaus, although without deeper analysis, this cannot be stated with certainty. While J-W mapping has better results in terms of the number of evaluations, parity mapping leads in accuracy. It can be speculated that if some orbitals were cleverly frozen in J-W mapping, it might provide the best results. However, the lowest energy now was achieved with parity mapping with reduction, at -1.1424689009110216 Hartree.

The last set examined is cc-PVDZ. Here, as already indicated, drastic limitations had to be applied to both the VQE simulation (and likewise the classical calculation) due to the comprehensiveness of the set, to make the calculations feasible at all. Using the *Active Space Transformer* function, the original 20 qubits were reduced to 8 closest to the Fermi level. The question of the optimality of this reduction and the very sensibleness of using this set under such restrictions is valid, but the aim of this section is to provide a high-level view of the simulation as a whole and not to delve into details. The resulting graph, captured in figure II.5, was created in the file `H2-PES-CCPVZD-PARITY (REDUCED) -WHOLE.ipynb`.

It is evident that VQE no longer provides satisfactory results here. While the classical calculation, even with necessary limitations, provides a quality estimate of the energy, VQE is unusable in the given setup. For this reason, more detailed graphs in the area of the minimum will not be presented here. Certainly, it would be possible to perform a smarter reduction of qubits or use a more suitable optimization technique, etc., but that is not the aim of this section.

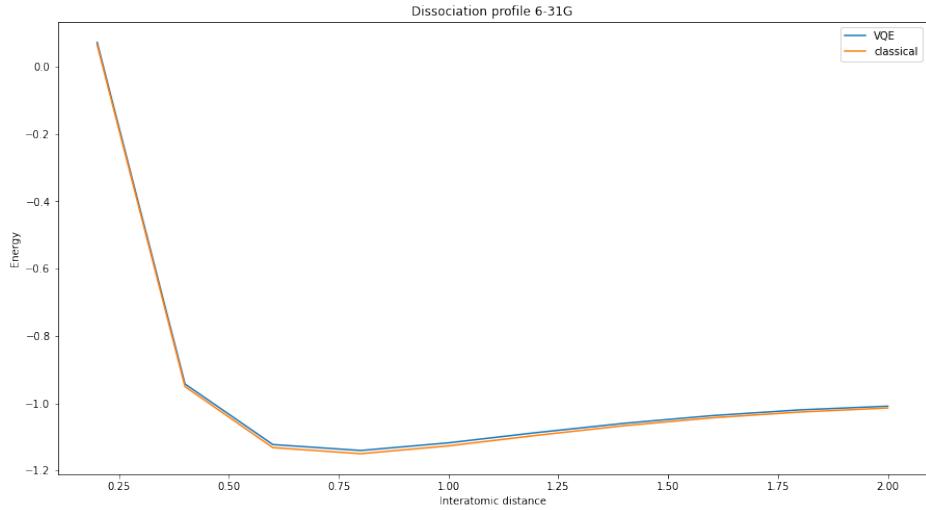


Figure II.3: Dissociation curve for molecular hydrogen when choosing the 6-31G basis set and parity mapping with two-qubit reduction. Samples counted at ten equidistant points within 0.2 to 2 Ångström. Number of cost function evaluations: 4089. Energy given in Hartree units and at the lowest point the VQE gives -1.1405381223642903 Hartree.

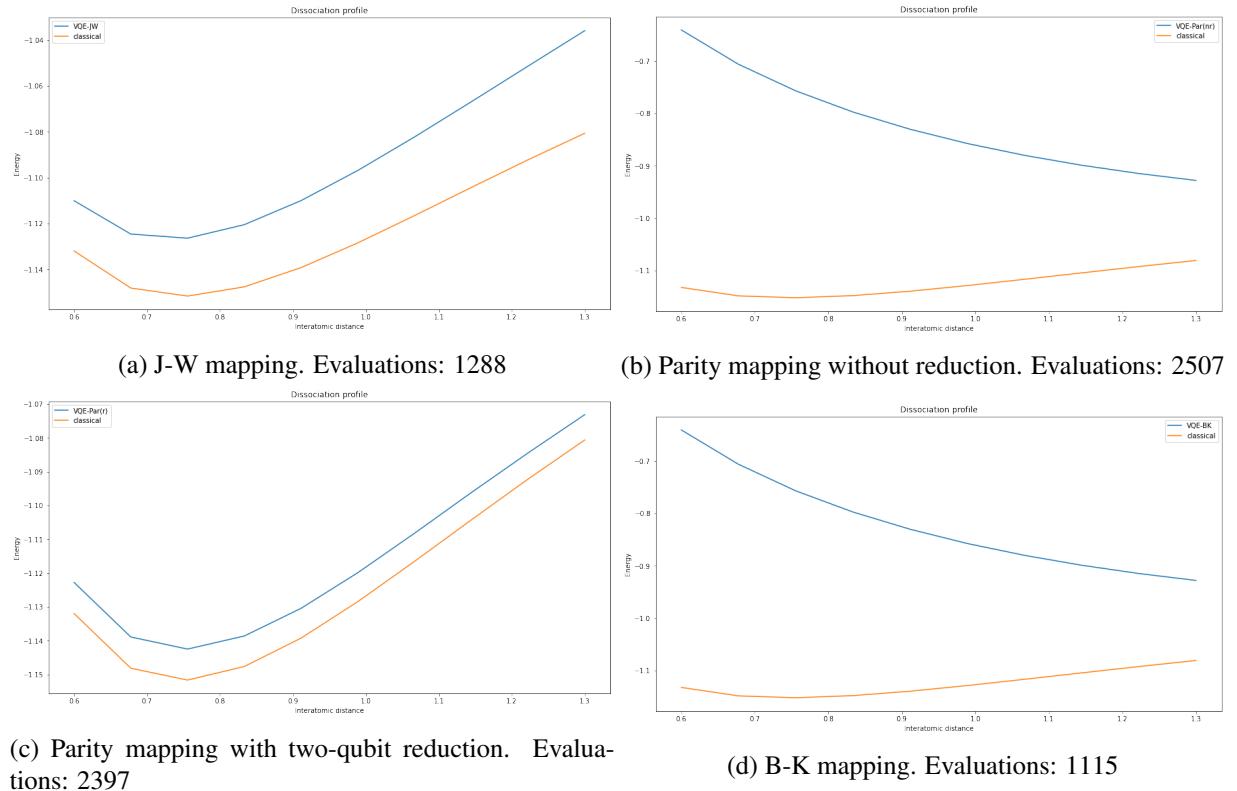


Figure II.4: Detail of the dissociation curves in the narrow range of 0.6 to 1.3 Ångström for the 6-31G basis set in different mappings. Samples counted at ten equidistant points.

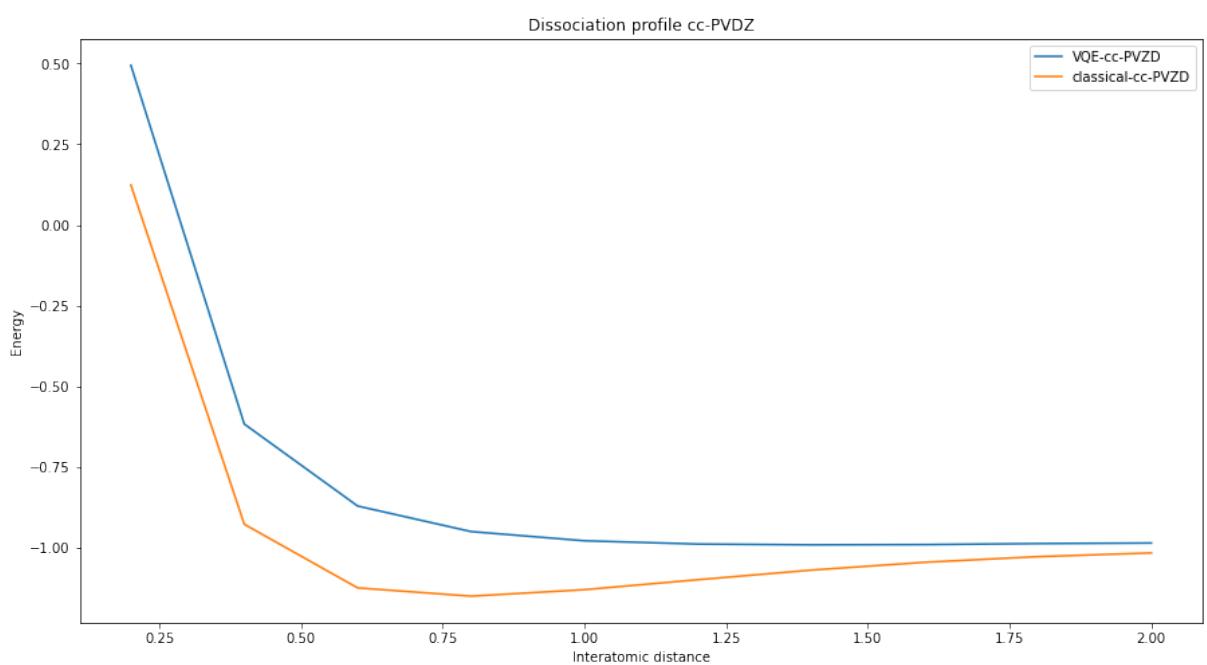


Figure II.5: Dissociation curve for molecular hydrogen when choosing the cc-PVDZ basis set and parity mapping with reduction to 8 qubits. Samples counted at ten equidistant points within 0.2 to 2 Ångström. Number of cost function evaluations: 3311. Energy given in Hartree units and at the lowest point the VQE gives -0.9906013889291293 Hartree.

II.4 Comparison of optimization techniques, ansatze and mappings

In this section, a more detailed comparison of VQE performance under different settings will be conducted. The STO-3G basis set will primarily be considered, but a condensed comparison will also be made for the 6-31G set. The molecule of hydrogen with an optimal nuclear distance of 0.735 Ångströms will be considered. For STO-3G, the compared ansätze will be: UCCSD, PUCCD, SUCCD (physically motivated), and TwoLocal (heuristic). The compared optimization techniques will be: Nelder-Mead algorithm, SPSA, L-BFGS-B, and COBYLA. Finally, the mappings will be the standard J-W, parity, and B-K. For 6-31G, the same optimization techniques will be considered, but only for J-W mapping, and among the ansätze, only UCCSD and TwoLocal will be considered.

Different settings will be evaluated in terms of: the number of qubits (always the same for the chosen basis set), the number of parameters forming the search space, the number of CNOT gates (because these contribute the most to error rates on real devices), and the number of cost function evaluations. Regarding the last criterion, it's worth mentioning that SPSA does not have well-defined termination criteria, so it always completes the maximum manually chosen number of iterations (500), leading to 1000 cost function evaluations, which may not be indicative.

The program conducting this comparison is written in the file `H2-COMPARISONS.ipynb`, which adds a row with results to the chosen CSV table for each evaluation under given settings and simultaneously generates a graph showing the dependency of energy computation on the number of iterations.

First, the STO-3G basis set will be considered. The output consists of three different tables, each for one mapping (J-W, parity, and B-K), with each row corresponding to one graph in the subsequent combined figure. The first table II.6 and its corresponding graphs II.7 pertain to J-W mapping. The second table II.8 and its corresponding graphs II.9 pertain to parity mapping. Finally, the table II.10 and its corresponding graphs II.11 pertain to B-K mapping. The first thing evident from the results

	optimizer	ansatz	rotation_blocks	entanglement_blocks	entanglement	repetitions	# of qubits	# of parameters	# of CNOTs	# of evaluations	error (mHa)
0	NELDER_MEAD	UCCSD	NaN	NaN	NaN	NaN	4	3	56	142	1.711387e-06
1	NELDER_MEAD	PUCCD	NaN	NaN	NaN	NaN	4	1	48	34	5.586240e-07
2	NELDER_MEAD	SUCCD	NaN	NaN	NaN	NaN	4	1	48	34	5.586231e-07
3	NELDER_MEAD	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	264	2.030704e+01
4	SPSA	UCCSD	NaN	NaN	NaN	NaN	4	3	56	1000	9.117104e-02
5	SPSA	PUCCD	NaN	NaN	NaN	NaN	4	1	48	1000	-5.107026e-12
6	SPSA	SUCCD	NaN	NaN	NaN	NaN	4	1	48	1000	-5.107026e-12
7	SPSA	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	1000	6.176793e+02
8	L_BFGS_B	UCCSD	NaN	NaN	NaN	NaN	4	3	56	24	9.521273e-10
9	L_BFGS_B	PUCCD	NaN	NaN	NaN	NaN	4	1	48	10	-4.662937e-12
10	L_BFGS_B	SUCCD	NaN	NaN	NaN	NaN	4	1	48	10	-4.662937e-12
11	L_BFGS_B	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	36	2.030704e+01
12	COBYLA	UCCSD	NaN	NaN	NaN	NaN	4	3	56	50	2.297850e-05
13	COBYLA	PUCCD	NaN	NaN	NaN	NaN	4	1	48	22	1.442381e-05
14	COBYLA	SUCCD	NaN	NaN	NaN	NaN	4	1	48	22	1.442381e-05
15	COBYLA	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	112	2.030704e+01

Figure II.6: Comparison table for STO-3G in J-W mapping.

is that physically motivated ansätze perform significantly better than heuristic ones. This is strongly determined by the small size of the basis set and the resulting number of qubits. For larger simulations on NISQ devices, this difference may be diminished. This is due, among other things, to the number

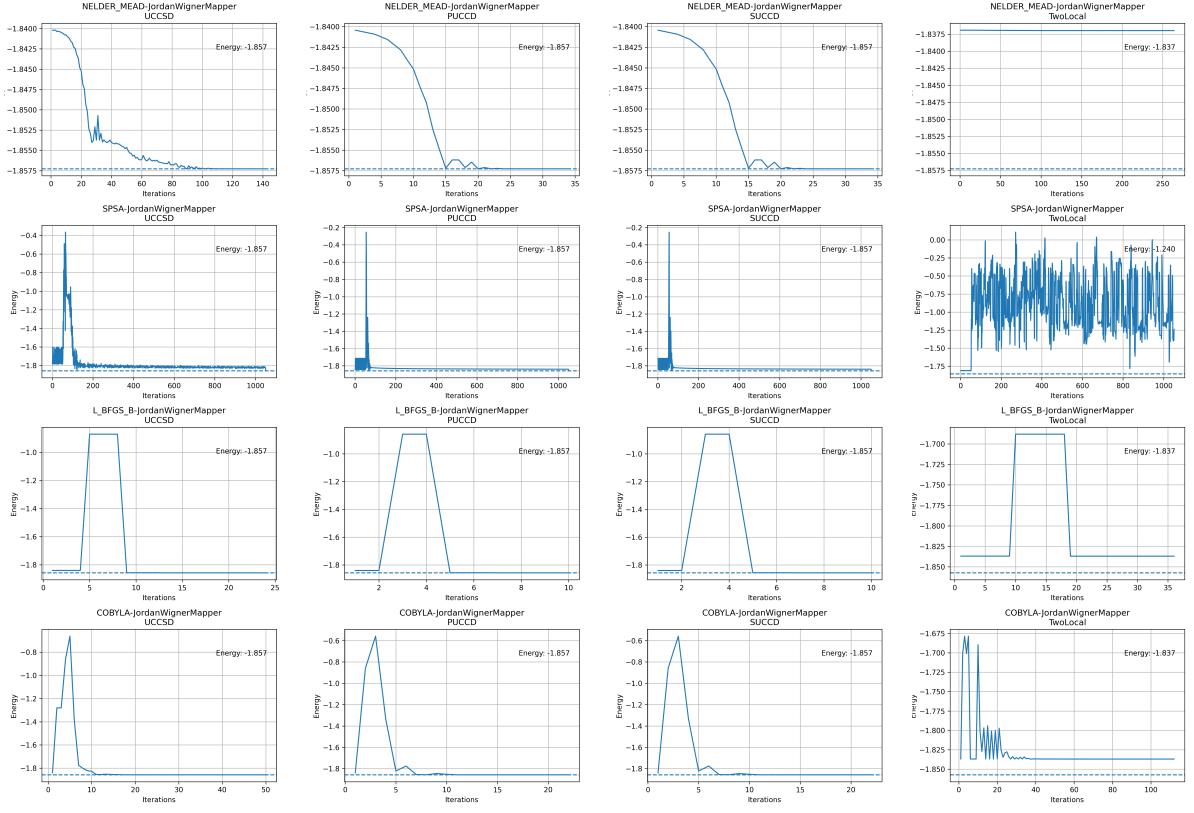


Figure II.7: Plots of the dependence of energy on the number of iterations for STO-3G in J-W mapping.

of CNOT gates, of which the heuristic ansatz uses much fewer. Another observation valid for all three mappings is that SPSA and L-BFGS-B provide the highest accuracy. However, even with the remaining two optimizers, achieving sufficient accuracy is not a problem. This is again due to the problem's size and is thus not very indicative. Overall, the PUCCD and SUCCD ansätze come out best from the comparison, and for this problem, there is only a slight difference between the mapping choices.

In some respects, the comparison for the 6-31G set is more interesting. For J-W mapping and ansätze limited only to UCCSD and TwoLocal, this is shown in table II.12 and the corresponding graphs II.13. There is still a visible (though smaller) advantage of physically motivated ansätze. However, the difference in the number of CNOT gates is now vast in favor of the heuristic ansatz. It is very interesting to note what happens when choosing SPSA optimization and the UCCSD ansatz. It is evident that the error remains large even with a large number of evaluations, and after several iterations, the energy stabilized at a certain value and did not improve further. This may indicate that the optimization got stuck in a local minimum of the parametric space.

	optimizer	ansatz	rotation_blocks	entanglement_blocks	entanglement	repetitions	# of qubits	# of parameters	# of CNOTs	# of evaluations	error (mHa)
0	NELDER_MEAD	UCCSD	NaN	NaN	NaN	NaN	4	3	38	142	1.711394e-06
1	NELDER_MEAD	PUCCD	NaN	NaN	NaN	NaN	4	1	32	34	5.586289e-07
2	NELDER_MEAD	SUCCD	NaN	NaN	NaN	NaN	4	1	32	34	5.586289e-07
3	NELDER_MEAD	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	273	2.030704e+01
4	SPSA	UCCSD	NaN	NaN	NaN	NaN	4	3	38	1000	3.581566e-02
5	SPSA	PUCCD	NaN	NaN	NaN	NaN	4	1	32	1000	-2.220446e-13
6	SPSA	SUCCD	NaN	NaN	NaN	NaN	4	1	32	1000	-2.220446e-13
7	SPSA	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	1000	6.009360e+02
8	L_BFGS_B	UCCSD	NaN	NaN	NaN	NaN	4	3	38	24	8.171241e-10
9	L_BFGS_B	PUCCD	NaN	NaN	NaN	NaN	4	1	32	10	6.217249e-12
10	L_BFGS_B	SUCCD	NaN	NaN	NaN	NaN	4	1	32	10	6.217249e-12
11	L_BFGS_B	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	45	2.030704e+01
12	COBYLA	UCCSD	NaN	NaN	NaN	NaN	4	3	38	49	6.008956e-06
13	COBYLA	PUCCD	NaN	NaN	NaN	NaN	4	1	32	22	6.246212e-06
14	COBYLA	SUCCD	NaN	NaN	NaN	NaN	4	1	32	22	6.246212e-06
15	COBYLA	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	128	2.030704e+01

Figure II.8: Comparison table for STO-3G in parity mapping.

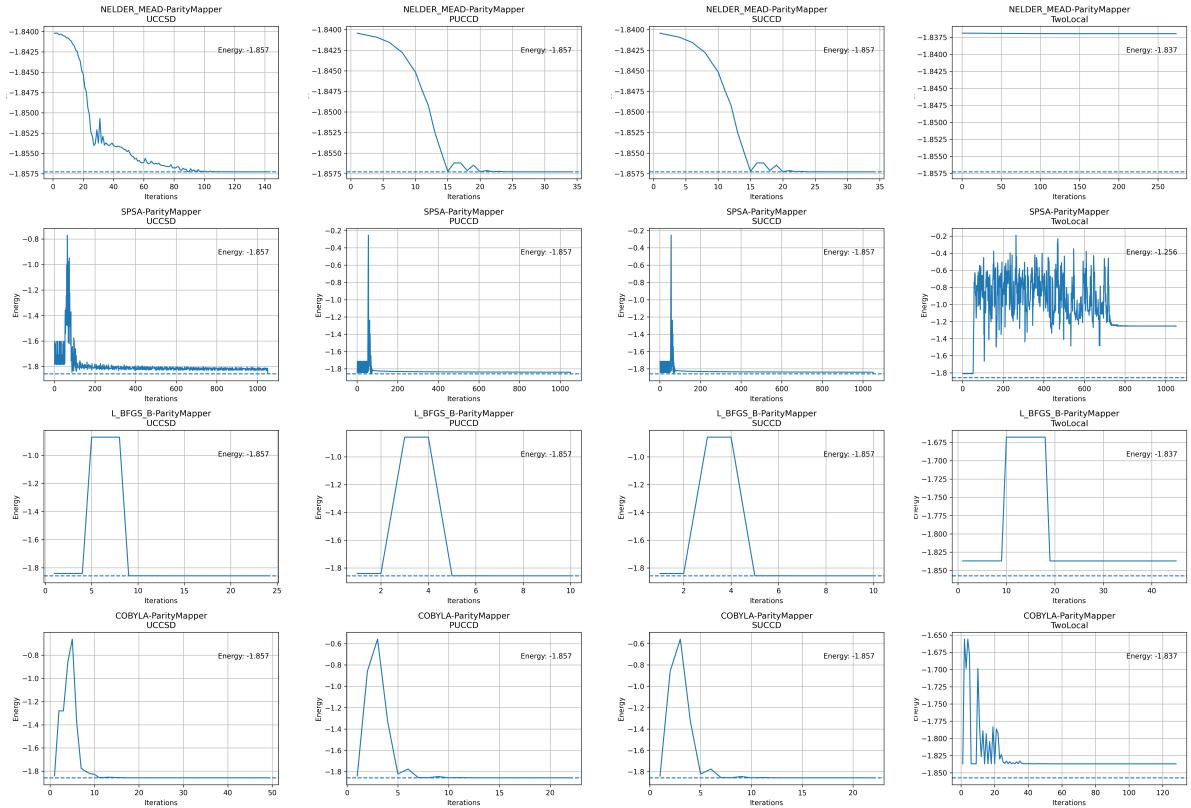


Figure II.9: Plots of the dependence of energy on the number of iterations for STO-3G in parity mapping.

	optimizer	ansatz	rotation blocks	entanglement_blocks	entanglement	repetitions	# of qubits	# of parameters	# of CNOTs	# of evaluations	error (mHa)
0	NELDER_MEAD	UCCSD	NaN		NaN	NaN	4	3	38	142	1.711393e-06
1	NELDER_MEAD	PUCCD	NaN		NaN	NaN	4	1	32	34	5.586276e-07
2	NELDER_MEAD	SUCCD	NaN		NaN	NaN	4	1	32	34	5.586276e-07
3	NELDER_MEAD	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	242	2.030704e+01
4	SPSA	UCCSD	NaN		NaN	NaN	4	3	38	1000	1.962271e-02
5	SPSA	PUCCD	NaN		NaN	NaN	4	1	32	1000	-4.440892e-13
6	SPSA	SUCCD	NaN		NaN	NaN	4	1	32	1000	-4.440892e-13
7	SPSA	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	1000	6.009360e+02
8	L_BFGS_B	UCCSD	NaN		NaN	NaN	4	3	38	24	6.450396e-10
9	L_BFGS_B	PUCCD	NaN		NaN	NaN	4	1	32	10	2.220446e-12
10	L_BFGS_B	SUCCD	NaN		NaN	NaN	4	1	32	10	2.220446e-12
11	L_BFGS_B	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	54	2.030704e+01
12	COBYLA	UCCSD	NaN		NaN	NaN	4	3	38	49	6.008953e-06
13	COBYLA	PUCCD	NaN		NaN	NaN	4	1	32	23	3.115048e-05
14	COBYLA	SUCCD	NaN		NaN	NaN	4	1	32	23	3.115048e-05
15	COBYLA	TwoLocal	['ry', 'rz']	cz	full	1.0	4	8	6	110	2.030704e+01

Figure II.10: Comparison table for STO-3G in B-K mapping.

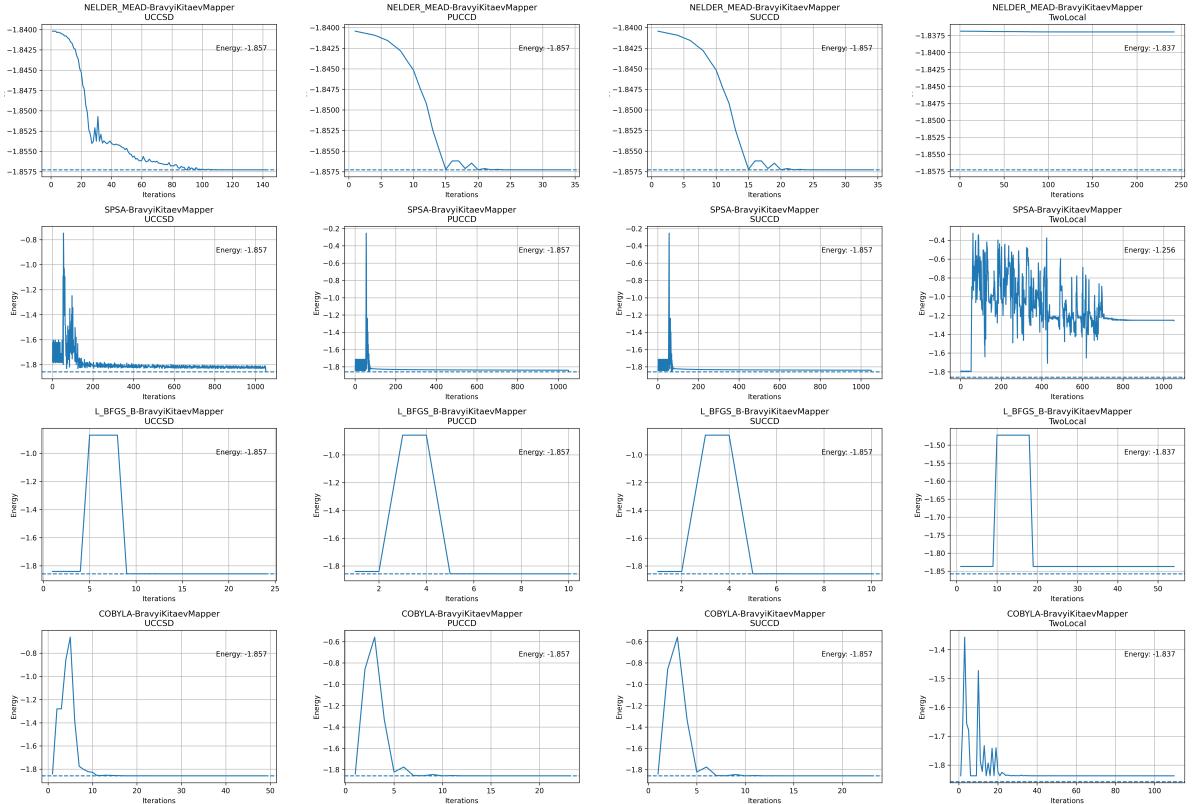


Figure II.11: Plots of the dependence of energy on the number of iterations for STO-3G in B-K mapping.

	optimizer	ansatz	rotation blocks	entanglement_blocks	entanglement	repetitions	# of qubits	# of parameters	# of CNOTs	# of evaluations	error (mHa)
0	NELDER_MEAD	UCCSD	NaN	NaN	NaN	NaN	8	15	768	628	1.824655e+00
1	NELDER_MEAD	TwoLocal	['ry', 'rz']	cz	full	1.0	8	16	28	645	2.480543e+01
2	SPSA	UCCSD	NaN	NaN	NaN	NaN	8	15	768	1000	4.920549e+02
3	SPSA	TwoLocal	['ry', 'rz']	cz	full	1.0	8	16	28	1000	1.816333e+03
4	L_BFGS_B	UCCSD	NaN	NaN	NaN	NaN	8	15	768	192	5.003775e-09
5	L_BFGS_B	TwoLocal	['ry', 'rz']	cz	full	1.0	8	16	28	153	2.480496e+01
6	COBYLA	UCCSD	NaN	NaN	NaN	NaN	8	15	768	266	3.829445e-05
7	COBYLA	TwoLocal	['ry', 'rz']	cz	full	1.0	8	16	28	268	2.480496e+01

Figure II.12: Comparison table for 6-31G in J-W mapping.

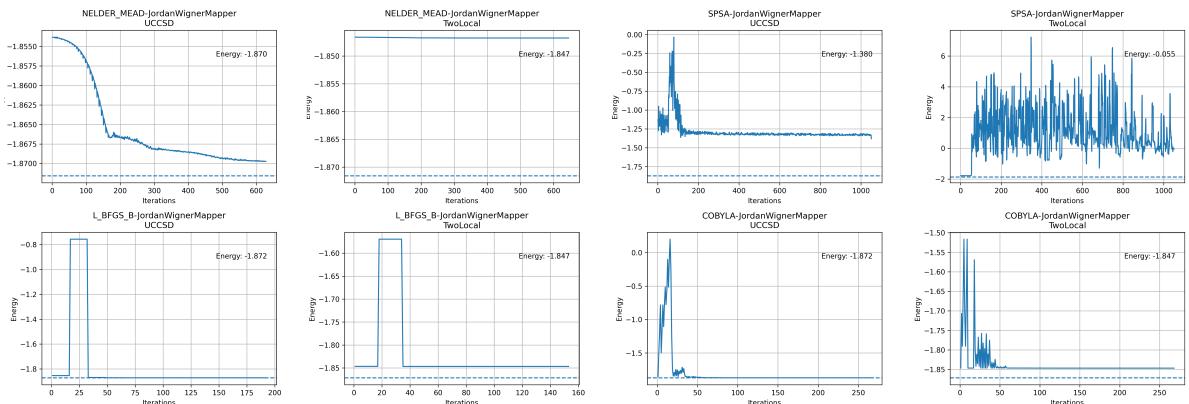


Figure II.13: Plots of the dependence of energy on the number of iterations for 6-31G in J-W mapping.

Conclusion

In this project, the most important methods of simulating quantum chemistry on quantum computers were summarized. In this direction, the research task can serve as an entry point into this issue for researchers involved in quantum informatics. In addition to classical computational chemistry methods, two directions of algorithm development for this area were discussed. This division was made based on the availability of devices capable of error correction.

In the long term, with the arrival of scalable error-correcting devices, methods based on the phase estimation algorithm seem to be the best option so far. This approach has guaranteed advantages over classical simulation, but the depth of circuits necessary for its implementation prevents its use without error correction. Conservative estimates speak of a 20-year horizon before such devices are available. It is therefore natural to look for ways to simulate quantum systems even on current or soon-to-be-available NISQ devices.

In this regard, VQE appears most promising, where long coherence times are exchanged for repeated state initializations and measurements. VQE for quantum chemistry simulation has been tested on several platforms, and sufficient accuracy has been achieved for very small molecules. However, as of the writing of this thesis, the so-called *quantum advantage*, where a quantum computer would outperform classical simulation of a molecule, has not yet been achieved. Optimistic estimates, however, suggest that this milestone could come within the next two to five years.

In both cases, research is in full swing, and the number of new publications is rapidly increasing. Especially with VQE, the algorithmic progress has been tremendous, and the complexity of simulation has been drastically reduced. It might seem that with the arrival of error-correcting quantum computers, the research conducted in the field of VQE and other techniques designed for NISQ devices would become obsolete, but the opposite is true. Publications already exist that aim to bridge these approaches and select the best from each of them [44], [66]. It is easy to imagine, for example, using VQE to better estimate the initial state for the subsequent phase estimation algorithm.

It needs to be said that essentially all considered techniques for simulating quantum chemistry have a strong foundation in classical computational techniques. However, it is not excluded that a purely quantum approach without any classical counterpart or foundation could be invented.

It is clear that achieving truly breakthrough results will require the collaboration of hardware engineers, quantum informatics experts, and quantum chemists. The connection and mutual sharing of ideas from these fields seem to be the right path towards simulations that are currently unparalleled. Nowadays, quantum simulations are mainly used to provide a theoretical basis for experimental results. In the future, with the development of quantum devices, it could happen that simulations might be more accurate than the experiment itself. This, for example, in cooperation with recent successes in machine learning in estimating protein structures, could bring absolutely groundbreaking results in the design of new compounds and materials. However, it is evident that further research in many areas is needed to realize these visions.

Literatura

- [1] Abrams, D. S. and Lloyd, S. “Simulation of many-body Fermi systems on a universal quantum computer.” In: *Physical Review Letters* 79.13 (1997), p. 2586.
- [2] ANIS, M. S. et al. *Qiskit: An Open-source Framework for Quantum Computing*. 2021. doi: [10.5281/zenodo.2573505](https://doi.org/10.5281/zenodo.2573505).
- [3] Aspuru-Guzik, A., Dutoi, A. D., Love, P. J., and Head-Gordon, M. “Simulated quantum computation of molecular energies.” In: *Science* 309.5741 (2005), pp. 1704–1707.
- [4] Babbush, R., Gidney, C., Berry, D. W., Wiebe, N., McClean, J., Paler, A., Fowler, A., and Neven, H. “Encoding electronic spectra in quantum circuits with linear T complexity.” In: *Physical Review X* 8.4 (2018), p. 041015.
- [5] Babbush, R., Wiebe, N., McClean, J., McClain, J., Neven, H., and Chan, G. K.-L. “Low-depth quantum simulation of materials.” In: *Physical Review X* 8.1 (2018), p. 011044.
- [6] Barkoutsos, P. K., Gonthier, J. F., Sokolov, I., Moll, N., Salis, G., Fuhrer, A., Ganzhorn, M., Egger, D. J., Troyer, M., Mezzacapo, A., et al. “Quantum algorithms for electronic structure calculations: Particle-hole Hamiltonian and optimized wave-function expansions.” In: *Physical Review A* 98.2 (2018), p. 022322.
- [7] Bartlett, R. J. and Musiał, M. “Coupled-cluster theory in quantum chemistry.” In: *Reviews of Modern Physics* 79.1 (2007), p. 291.
- [8] Berry, D. W. and Childs, A. M. “Black-box Hamiltonian simulation and unitary implementation.” In: *arXiv preprint arXiv:0910.4157* (2009).
- [9] Berry, D. W., Childs, A. M., Cleve, R., Kothari, R., and Somma, R. D. “Simulating Hamiltonian dynamics with a truncated Taylor series.” In: *Physical review letters* 114.9 (2015), p. 090502.
- [10] Berry, D. W., Kieferová, M., Scherer, A., Sanders, Y. R., Low, G. H., Wiebe, N., Gidney, C., and Babbush, R. “Improved techniques for preparing eigenstates of fermionic Hamiltonians.” In: *npj Quantum Information* 4.1 (2018), pp. 1–7.
- [11] Born, M. and Oppenheimer, R. “Zur quantentheorie der moleküle.” In: *Annalen der physik* 389.20 (1927), pp. 457–484.
- [12] Brown, K. R., Kim, J., and Monroe, C. “Co-designing a scalable quantum computer with trapped atomic ions.” In: *npj Quantum Information* 2.1 (2016), pp. 1–10.
- [13] Broyden, C. G. “The convergence of a class of double-rank minimization algorithms 1. general considerations.” In: *IMA Journal of Applied Mathematics* 6.1 (1970), pp. 76–90.
- [14] Byrd, R. H., Lu, P., Nocedal, J., and Zhu, C. “A limited memory algorithm for bound constrained optimization.” In: *SIAM Journal on scientific computing* 16.5 (1995), pp. 1190–1208.

- [15] Cao, Y., Romero, J., Olson, J. P., Degroote, M., Johnson, P. D., Kieferová, M., Kivlichan, I. D., Menke, T., Peropadre, B., Sawaya, N. P., et al. “Quantum chemistry in the age of quantum computing.” In: *Chemical reviews* 119.19 (2019), pp. 10856–10915.
- [16] Childs, A. M. “On the relationship between continuous-and discrete-time quantum walk.” In: *Communications in Mathematical Physics* 294.2 (2010), pp. 581–603.
- [17] Childs, A. M., Ostrander, A., and Su, Y. “Faster quantum simulation by randomization.” In: *Quantum* 3 (2019), p. 182.
- [18] Childs, A. M. and Wiebe, N. “Hamiltonian simulation using linear combinations of unitary operations.” In: *arXiv preprint arXiv:1202.5822* (2012).
- [19] Dallaire-Demers, P.-L., Romero, J., Veis, L., Sim, S., and Aspuru-Guzik, A. “Low-depth circuit ansatz for preparing correlated fermionic states on a quantum computer.” In: *Quantum Science and Technology* 4.4 (2019), p. 045005.
- [20] Devoret, M. H., Wallraff, A., and Martinis, J. M. “Superconducting qubits: A short review.” In: *arXiv preprint cond-mat/0411174* (2004).
- [21] Evans, M. G. and Polanyi, M. “Some applications of the transition state method to the calculation of reaction velocities, especially in solution.” In: *Transactions of the Faraday Society* 31 (1935), pp. 875–894.
- [22] Eyring, H. “The activated complex in chemical reactions.” In: *The Journal of Chemical Physics* 3.2 (1935), pp. 107–115.
- [23] Ganzhorn, M., Egger, D. J., Barkoutsos, P., Ollitrault, P., Salis, G., Moll, N., Roth, M., Fuhrer, A., Mueller, P., Woerner, S., et al. “Gate-efficient simulation of molecular eigenstates on a quantum computer.” In: *Physical Review Applied* 11.4 (2019), p. 044092.
- [24] Hammersley, J. and Morton, K. “A new Monte Carlo technique: antithetic variates.” In: *Mathematical proceedings of the Cambridge philosophical society*. Vol. 52. 3. Cambridge University Press. 1956, pp. 449–475.
- [25] Hartree, D. R. “The wave mechanics of an atom with a non-Coulomb central field. Part I. Theory and methods.” In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 24. 1. Cambridge university press. 1928, pp. 89–110.
- [26] Havlíček, V., Troyer, M., and Whitfield, J. D. “Operator locality in the quantum simulation of fermionic models.” In: *Physical Review A* 95.3 (2017), p. 032332.
- [27] Helgaker, T., Jorgensen, P., and Olsen, J. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.
- [28] Hempel, C., Maier, C., Romero, J., McClean, J., Monz, T., Shen, H., Jurcevic, P., Lanyon, B. P., Love, P., Babbush, R., et al. “Quantum chemistry calculations on a trapped-ion quantum simulator.” In: *Physical Review X* 8.3 (2018), p. 031022.
- [29] Jordan, P. and Wigner, E. P. “Über das paulische äquivalenzverbot.” In: *The Collected Works of Eugene Paul Wigner*. Springer, 1993, pp. 109–129.
- [30] Kandala, A., Mezzacapo, A., Temme, K., Takita, M., Brink, M., Chow, J. M., and Gambetta, J. M. “Hardware-efficient variational quantum eigensolver for small molecules and quantum magnets.” In: *Nature* 549.7671 (2017), pp. 242–246.
- [31] Kohn, W. “Nobel Lecture: Electronic structure of matter—wave functions and density functionals.” In: *Reviews of Modern Physics* 71.5 (1999), p. 1253.

- [32] Kolda, T. G., Lewis, R. M., and Torczon, V. “Optimization by direct search: New perspectives on some classical and modern methods.” In: *SIAM review* 45.3 (2003), pp. 385–482.
- [33] Lahtinen, V. and Pachos, J. “A short introduction to topological quantum computation.” In: *SciPost Physics* 3.3 (2017), p. 021.
- [34] Low, G. H. and Chuang, I. L. “Hamiltonian simulation by qubitization.” In: *Quantum* 3 (2019), p. 163.
- [35] Low, G. H. and Chuang, I. L. “Optimal Hamiltonian simulation by quantum signal processing.” In: *Physical review letters* 118.1 (2017), p. 010501.
- [36] Martin, R. M. *Electronic structure: basic theory and practical methods*. Cambridge university press, 2020.
- [37] McArdle, S., Endo, S., Aspuru-Guzik, A., Benjamin, S. C., and Yuan, X. “Quantum computational chemistry.” In: *Reviews of Modern Physics* 92.1 (2020), p. 015003.
- [38] McClean, J. R., Babbush, R., Love, P. J., and Aspuru-Guzik, A. “Exploiting locality in quantum computation for quantum chemistry.” In: *The journal of physical chemistry letters* 5.24 (2014), pp. 4368–4380.
- [39] McClean, J. R., Boixo, S., Smelyanskiy, V. N., Babbush, R., and Neven, H. “Barren plateaus in quantum neural network training landscapes.” In: *Nature communications* 9.1 (2018), pp. 1–6.
- [40] McClean, J. R., Romero, J., Babbush, R., and Aspuru-Guzik, A. “The theory of variational hybrid quantum-classical algorithms.” In: *New Journal of Physics* 18.2 (2016), p. 023023.
- [41] Nelder, J. A. and Mead, R. “A simplex method for function minimization.” In: *The computer journal* 7.4 (1965), pp. 308–313.
- [42] Nielsen, M. A. and Chuang, I. *Quantum computation and quantum information*. 2002.
- [43] Oliveira, I., Sarthour Jr, R., Bonagamba, T., Azevedo, E., and Freitas, J. C. *NMR quantum information processing*. Elsevier, 2011.
- [44] O’Brien, T., Tarasinski, B., and Terhal, B. “Quantum phase estimation for noisy, small-scale experiments.” In: *arXiv preprint arXiv:1809.09697* (2018).
- [45] Patterson, J. D. and Bailey, B. C. *Solid-state physics: introduction to the theory*. Springer Science & Business Media, 2007.
- [46] Peruzzo, A et al. “A variational eigenvalue solver on a quantum processor. eprint.” In: *arXiv preprint arXiv:1304.3061* (2013).
- [47] Peruzzo, A., McClean, J., Shadbolt, P., Yung, M.-H., Zhou, X.-Q., Love, P. J., Aspuru-Guzik, A., and O’brien, J. L. “A variational eigenvalue solver on a photonic quantum processor.” In: *Nature communications* 5.1 (2014), pp. 1–7.
- [48] Poulin, D., Kitaev, A., Steiger, D. S., Hastings, M. B., and Troyer, M. “Quantum algorithm for spectral measurement with a lower gate count.” In: *Physical review letters* 121.1 (2018), p. 010501.
- [49] Pyykkö, P. “Relativistic Effects in Chemistry: More Common Than You Thought.” In: *Annual Review of Physical Chemistry* 63.1 (2012). PMID: 22404585, pp. 45–64. doi: 10.1146/annurev-physchem-032511-143755. URL: <https://doi.org/10.1146/annurev-physchem-032511-143755>.
- [50] Quantum, G. A., Collaborators*†, Arute, F., Arya, K., Babbush, R., Bacon, D., Bardin, J. C., Barends, R., Boixo, S., Broughton, M., Buckley, B. B., et al. “Hartree-Fock on a superconducting qubit quantum computer.” In: *Science* 369.6507 (2020), pp. 1084–1089.

- [51] Romero, J., Babbush, R., McClean, J. R., Hempel, C., Love, P. J., and Aspuru-Guzik, A. “Strategies for quantum computing molecular energies using the unitary coupled cluster ansatz.” In: *Quantum Science and Technology* 4.1 (2018), p. 014008.
- [52] Romero, J., Olson, J. P., and Aspuru-Guzik, A. “Quantum autoencoders for efficient compression of quantum data.” In: *Quantum Science and Technology* 2.4 (2017), p. 045001.
- [53] Roothaan, C. C. J. “New developments in molecular orbital theory.” In: *Reviews of modern physics* 23.2 (1951), p. 69.
- [54] Rubin, N. C., Babbush, R., and McClean, J. “Application of fermionic marginal constraints to hybrid quantum algorithms.” In: *New Journal of Physics* 20.5 (2018), p. 053020.
- [55] Seeley, J. T., Richard, M. J., and Love, P. J. “The Bravyi-Kitaev transformation for quantum computation of electronic structure.” In: *The Journal of chemical physics* 137.22 (2012), p. 224109.
- [56] Shen, Y., Zhang, X., Zhang, S., Zhang, J.-N., Yung, M.-H., and Kim, K. “Quantum implementation of the unitary coupled cluster for simulating molecular electronic structure.” In: *Physical Review A* 95.2 (2017), p. 020501.
- [57] Shikano, Y., Watanabe, H. C., Nakanishi, K. M., and Ohnishi, Y.-y. “Post-Hartree–Fock method in quantum chemistry for quantum computer.” In: *The European Physical Journal Special Topics* 230.4 (2021), pp. 1037–1051.
- [58] Somma, R., Ortiz, G., Gubernatis, J. E., Knill, E., and Laflamme, R. “Simulating physical phenomena by quantum networks.” In: *Physical Review A* 65.4 (2002), p. 042323.
- [59] Spall, J. C. “An overview of the simultaneous perturbation method for efficient optimization.” In: *Johns Hopkins apl technical digest* 19.4 (1998), pp. 482–492.
- [60] Suzuki, M. “General theory of fractal path integrals with applications to many-body theories and statistical physics.” In: *Journal of Mathematical Physics* 32.2 (1991), pp. 400–407.
- [61] Taube, A. G. and Bartlett, R. J. “New perspectives on unitary coupled-cluster theory.” In: *International journal of quantum chemistry* 106.15 (2006), pp. 3393–3401.
- [62] Tilly, J., Chen, H., Cao, S., Picozzi, D., Setia, K., Li, Y., Grant, E., Wossnig, L., Rungger, I., Booth, G. H., et al. “The variational quantum eigensolver: a review of methods and best practices.” In: *arXiv preprint arXiv:2111.05176* (2021).
- [63] Trotter, H. F. “On the product of semi-groups of operators.” In: *Proceedings of the American Mathematical Society* 10.4 (1959), pp. 545–551.
- [64] Truhlar, D. G., Garrett, B. C., and Klippenstein, S. J. “Current status of transition-state theory.” In: *The Journal of physical chemistry* 100.31 (1996), pp. 12771–12800.
- [65] Turney, J. M., Simmonett, A. C., Parrish, R. M., G, H. E., Evangelista, F., Fermann, J. T., Mintz, B. J., Burns, L. A., Wilke, J. J., Abrams, M. L., Russ, N. J., Leininger, M. L., Janssen, C. L., Seidl, E. T., Allen, W. D., Schaefer, H. F., King, R. A., Valeev, E. F., Sherrill, C. D., and Crawford, T. D. “Psi4: An open-source ab initio electronic structure program.” In: *WIREs Computational Molecular Science* 2 (2012), pp. 556–565. doi: 10.1002/wcms.93.
- [66] Wang, D., Higgott, O., and Brierley, S. “Accelerated variational quantum eigensolver.” In: *Physical review letters* 122.14 (2019), p. 140504.
- [67] Wecker, D., Bauer, B., Clark, B. K., Hastings, M. B., and Troyer, M. “Gate-count estimates for performing quantum chemistry on small quantum computers.” In: *Physical Review A* 90.2 (2014), p. 022305.

- [68] Wecker, D., Hastings, M. B., and Troyer, M. “Progress towards practical quantum variational algorithms.” In: *Physical Review A* 92.4 (2015), p. 042303.
- [69] Whitfield, J. D., Biamonte, J., and Aspuru-Guzik, A. “Simulation of electronic structure Hamiltonians using quantum computers.” In: *Molecular Physics* 109.5 (2011), pp. 735–750.
- [70] Yung, M.-H., Casanova, J., Mezzacapo, A., McClean, J., Lamata, L., Aspuru-Guzik, A., and Solano, E. “From transistor to trapped-ion computers for quantum chemistry.” In: *Scientific reports* 4.1 (2014), pp. 1–7.