# Università degli Studi di Padova

Department of Physics and Astronomy



# Excited States Forces With BSE


Relatore:

**Prof. Paolo Umari**

Candidato:

**Marah Alrahamneh**


**Anno Accademico 2023/2024**

# Abstract

Photochemical processes are fundamental to wide range of physical and chemical phenomena. Many of these processes involve structural changes, e.g. fragmatization, formation or elimination of defects. In order to understand these photo-induced structural changes using first principles calculations, we need to calculate the interatomic forces in the excited state accurately and efficiently. By combining the many-body Green's function (GW) and the Bethe-Salpeter equation formalism. We present a practical method for computing the excited state forces, which is implemented with the help of the Quantum espresso package and the GWL code. In order to test this approach we used the CO molecule, the results proved to be in agreement with the experiment.

# Contents

# List of Figures

**Title** **Page No.**

# Chapter One

# Bird's Eye View: Setting the Stage

> *The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved.*

Paul Dirac, *1929*

## 1.1 Introduction

In the vibrant world of material science, computer-based modeling and simulation have emerged as game-changers. They're not just tools for understanding and developing a wide array of materials; they're also crystal balls that help us foresee entirely new material classes. The real magic of these simulations lies in the fact that they let us boldly go where no lab has gone before, exploring materials in extreme conditions that would be off-limits in traditional lab setting. This ' no sample needed ' approach has led many top-tier labs to embrace a " **Computation first, then experiment** " approach. Generally, when we delve into the realm of computer simulations, as outlined in [1], we find two distinct paths:

1. **The Top-Down Approach:** which refers to empirical or semi-empirical methods. In this approach laws are crafted and tweaked to fit the experimental data. It is like reverse-engineering the laws of nature from what we observe, fine-tuning parameters extracted from experimental findings to get the perfect fit.

2. **The Bottom-Up Approach:** Here lies the domain of ab-initio, or " First principles " methods. True to its name, which literally translates to from the beginning, this approach doesn't rely on experimental data. Instead, it draws from the deep well of quantum mechanics, which offers a general, flexible and robust foundation for building our understanding from the ground up.

In this thesis, our focus will be predominantly on the bottom-up approach, to bridge the microscopic quantum world with the macroscopic experimental observations. The cornerstone of this bridge is Density functional theory. However, it is important to bear in mind that DFT, at its core, is a ground state theory, which makes it less effective for applications involving excited states. For example, a key component in material characterization are spectroscopic measurements [2]. These measurements

perturb the system under study, elevating it to an excited state. In this case one needs to look beyond DFT, mainly GW approximation and the Bethe-Salpeter equation.

Fig.1.1 shows a schematic representation of electronic band structure calculations using different theoretical methods, starting from Density Functional Theory to many-body perturbation theory methods like GW approximation and the Bethe-Salpeter Equation. Each theory picks up where the previous one naturally fails.

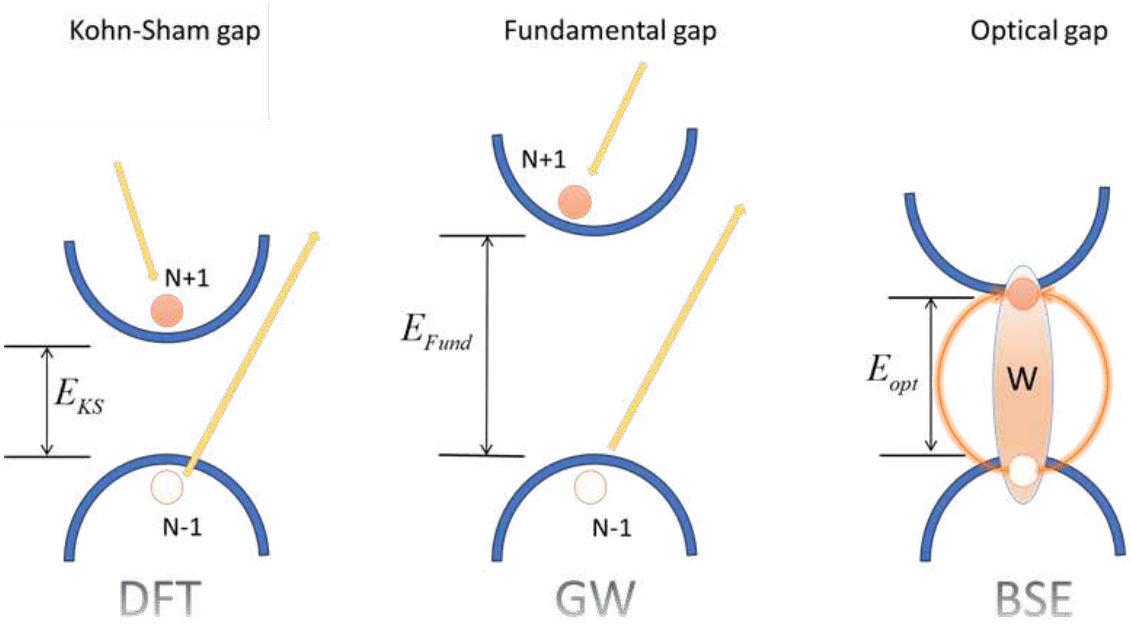



**Figure 1.1:** The electronic structure calculation hierarchy: from Density Functional Theory (DFT) to many-body perturbation theory methods like GW approximation and the Bethe-Salpeter Equation (BSE)

## 1.2 Theoretical Spectroscopy

Theoretical spectroscopy is one of the most active research fields in condensed matter physics [3]. Despite it's immense success in describing ground state properties, when it comes to theoretical spectroscopy Density functional theory falls short. It gives wrong quantitative results for charged excitation, for example, the whole debacle of the band-gap problem. This is where, one looks for a better description, which

is usually the case in ab initio modelling; there is no one size fit all solution. An effective approach to describe charged excitation can be obtained using a quasiparticle description. It predicts band gaps in good agreements with experiment [4]. However, it fails when it comes to neutral excitation and predicting the optical band gap for example. Neutral excitation require special treatment, which can be done using the Bethe-Salpeter Equation.

### 1.2.1 Photoemssion

In photoemssion experiments a photon with energy $h\nu$ irradiates a sample with energy $E(N,0)$, as a result an electron is ejected with a kinetic energy $E_{kin}$, as illustrated in Fig.1.2 [5]. Energy conservation dictates that

$$h\nu + E(N,0) = E(N-1,i) + E_{\text{kin}} \tag{1.1}$$

The ionization potential (IP), which is the energy required to remove an electron from the neutral many-body system with energy E(N,0), is defined as:

$$IP_i = E(N-1,i) - E(N,0) = h\nu - E_{\text{kin}} \tag{1.2}$$

Fig.1.2 also shows another process connected to photoemssion, i.e the inverse photoemession (IPES). An incident electron with energy $E_{kin}$ is scattered in a sample with energy $E(N,0)$ emitting bresmsstrahlng. Eventually, it goes down to lower energy level emitting a photon as a result with energy $h\nu$, the sample now contains an extra electron. From energy conservation:

$$E_{\text{kin}} + E(N,0) = E(N+1,i) + h\nu \tag{1.3}$$

The energy required to remove an electron from a negatively charge sample is known as the electron Affinity (EA) which can expressed as follows:

$$EA_i = E(N,0) - E(N+1,i) = h\nu - E_{\text{kin}} \tag{1.4}$$

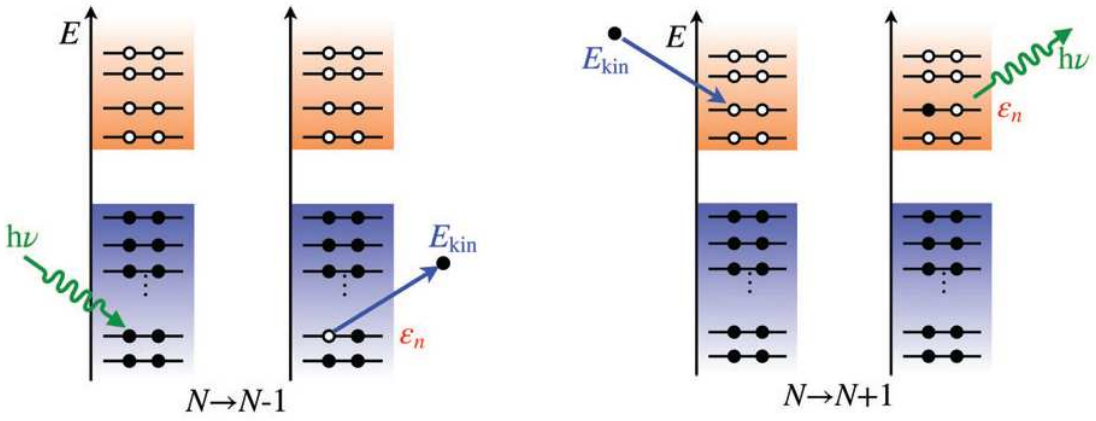



**Figure 1.2:** Illustration of the direct (left) and inverse (right) photoemission in an independent particle picture [6]. The wiggly green arrow indicates a photon entering (leaving) the sample in photoemssion (inverse photoemssion). When the photon excite the sample, an electron (black circle) is ejected leaving behind a hole (the white circle)

In our discussion thus far, we described photoemssion and inverse photoemssion in the context of the independent particle picture, where the two process complement each other. One is used to probe occupied states while the other the unoccupied ones as illustrated in Fig.1.2. This framework offers a simplified and straightforward model. In reality however, that's not the case. Experimantalist don't measure spectra with delta-function-like peaks as one would expect from the independent particle picture, rather the structure is more nuanced and rich with details. To interpret these spectra, one needs to go beyond the independent particle picture and delve into the realm of many-body interactions.

The many-body effects manifest themselves as a shift and re-normalization of the independent particle peaks. For instance, in photoemission a more complete picture

would be to consider the impact of the newly created hole on the rest of the sample. The presence of the hole creates a polarization cloud which dynamically screens the hole's charge. Thus, for a more accurate depiction one needs to abandon the independent picture of the charges and consider the "dressed" electrons or holes, i.e electrons or holes and their polarization cloud. These dressed charges are better known as quasiparticles. The quasiparticle picture offers a good interpretation of the experimental spectra. Later on we describe the GW approximation, which is the theoretical framework devised to analyze these charged excitation.

### 1.2.2 Optical Absorption

Contrary to photoemission where an electron is ejected from the material, in optical absorption, the electron upon absorbing the photon gets promoted to a higher energy level in the same sample leaving behind a positive vacancy, better known as a hole. The electron, now in an excited state, and the hole it leaves behind are not independent, they interact through attractive Coulomb interaction which leads to the formation of a bound pair, known as an exciton. The study of these excitons is crucial to understand the electronic response of materials to optical stimulation, determine their optical properties not just for absorption but emission as well through fluorescence or phosphorescence [7]. They also play a pivotal role in chemical reactions [8] and in driving biological processes [9]. This plays a pivotal role in practical applications, notably in photovoltaic technology and the advancement of optoelectronic devices, where the manipulation and control of such excitons are key [10, 11, 12].

### 1.2.3 Why BSE for neutral excitation?

Optical spectra is often used to explore and characterize materials properties, it captures qualities such as absorption, reflectivity, photoluminescence, and other optical behaviors. Hence why a reliable and efficient ab initio description of optical spectra is crucial to determine whether a material is suitable for a variety of technological applications, ranging from light-emitting devices and laser technology to photovoltaics [13].

While GW approximation manages to provide insights beyond what DFT can offer, particularly in describing charged excitation. Yet, it is not as effective in accurately evaluating optical spectra or charge-neutral excitation. This is where the Bethe-Salpeter equation comes into play. It goes beyond the independent quasi-particle approach, by adopting an effective two-body one, which accurately captures neutral excitation where electron-hole interactions are crucial [14]. This offers a more complete picture compared to GW approximation.

## 1.3 Excited States Forces

Neutrally excited systems undergoes structural changes, which can lead to a variety of subsequent processes such as photochemical reactions and energy transfer [15]. Fig.1.3 illustrates the potential energy curves for a molecular system, the solid represents the ground state energy ($E_0$) as a function of the bond length (R), and the dashed line represents the excited state energy ($E_s$). The shift in the equilibrium position on the excited state curve, relative to the ground state, signifies the structural changes, i.e changes in bond length in this simplified model, that occur upon excitation.

For ground state electronic structure calculations geometry optimization is one of the most preformed tasks [16]. So much so that most of the major electronic structure

packages have a selection of geometry optimization algorithms already implemented. However, when it comes to excited electronic states, this task becomes somewhat of a challenge [17].

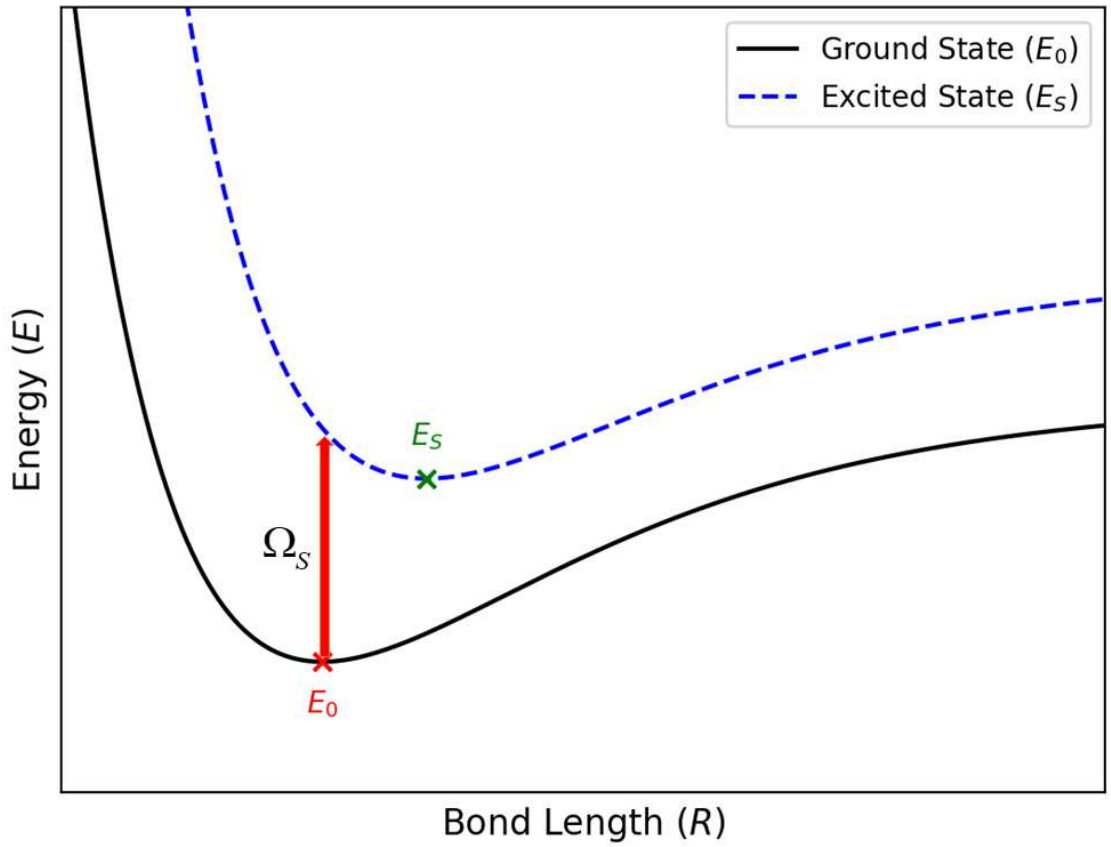



**Figure 1.3:** Schematic illustration of potential energy curves for ground and excited States

In general, and for more complicated potential energy surfaces, in the vicinity of the equilibrium geometry $\mathbf{x_0}$ the potential energy surface can be approximated as follows [18]:

$$E(\mathbf{x}) = E(\mathbf{x_0}) + g_0^T \Delta\mathbf{x} + \frac{1}{2}\Delta\mathbf{x}^T H_0 \Delta\mathbf{x} \tag{1.5}$$

where $\Delta\mathbf{x} = \mathbf{x} - \mathbf{x}_0$, $g_0$ is the gradient of the energy $\frac{dE}{d\mathbf{x}}$ at $\mathbf{x_0}$ which is the negative of the vector of the forces acting on the nuclei (I) ($F_I \equiv -\frac{dE}{d\mathbf{x}_I}$), since it is taken at the equilibrium position $\mathbf{x_0}$ it is equal to zero. $H_0$ is the Hessian matrix ($\frac{d^2E}{dx^2}$) also

known as the force constant matrix. At equilibrium all the eigenvalues of the Hessian matrix must be positive.

Thus far, we have outlined the foundational concepts without dwelling on the intricate details. To understand both the ground and excited states properties, one needs to work with quantum systems which consist of a large number of interacting electrons and nuclei. The challenges and intricacies of these systems escalate dramatically, giving rise to what is known in physics as the " **Many-Body Problem**".

## 1.4 The many-body problem

The many-body problem is a topic that is widely discussed in all branches of physics. But what is the criteria for a problem to be described as a Many-body problem? There are two essential ingredients: The obvious one being, there have to be many-bodies present and these bodies **have** to interact with each other. These interactions are critical for the many-body problem, as many phenomena observed in nature are emergent ones that cannot be explained solely by the fundamental laws governing individual particles. ***More is different*** as Anderson famously stated [19]. These interactions are so important that the many-body problem can be defined as " *The study of the effects of interaction between bodies on the behaviour of a many-body system* " [20].

## 1.5 Thesis Outline

The thesis is organized into five chapters, as follows:

- **Chapter 1: A Bird's Eye View: Setting the Stage**

  As the title suggests, this chapter serves as a prelude. It sets the tone for further discussions in subsequent chapters.

- **Chapter 2: Ground State properties**

  This chapter gives an overview of Density Functional Theory (DFT), it also discusses its shortcomings in examining excited states and the need to go beyond DFT.

- **Chapter 3: The Realm of Excited States**

  This chapter addresses the use of GW approximation to study charged excitation. It then describes the Bethe-Salpeter Equation (BSE) within the Tamm-Dancoff approximation to study neutral excitation.

- **Chapter 4: Excited State Forces and Structural Optimization**

  This chapter focuses on the development of a novel approach used for calculating excited state forces for a neutrally excited system in a perturbed Tamm-Dancoff manifold. It also discuss structural optimization using BFGS approach.

- **Chapter 5: From Theory to Practice: CO Molecule as a Test Run**
  The final chapter offers practical application of the theoretical concepts by using CO molecule as a test subject.

# Chapter Two

# Ground State properties

## 2.1 Introduction

A key goal of electronic structure calculations is to use quantum mechanic's first principles to study general molecular and solid-state systems, predict their properties and discover new materials, bridging theory and experiment.

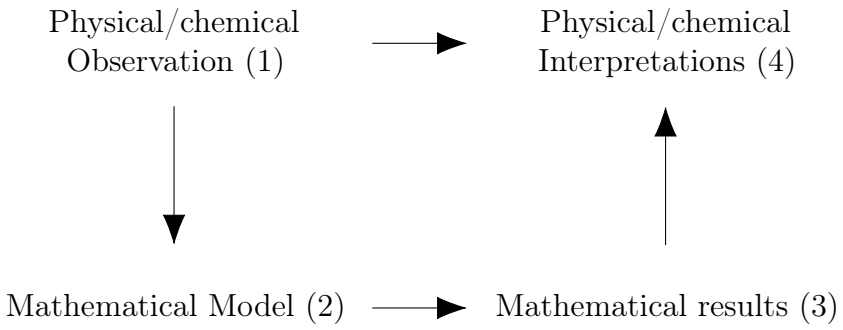



**Figure 2.1:** Flowchart illustrating the process of how mathematical modeling, physical predictions and observations are all interconnected [21].

In this chapter we want to give an overview of the formalism needed to describe the ground state properties. In theory, this can be done by providing a reliable and computationally accessible solutions to the non-relativistic time-independent Schrodinger equation [22]:

$$\hat{H}\Psi(\vec{r},\vec{R}) = E\Psi(\vec{r},\vec{R}) \tag{2.1}$$

where $\hat{H}$ is the non-relativistic Hamiltonian that governs the dynamics of the mechanical system of N electrons and M nuclei. $\Psi(\vec{r},\vec{R})$ is the many-body wave function which depends on both the electrons and nuclei positions in the many-body system.

$$H(\vec{r},\vec{R}) = -\frac{1}{2}\sum_{J}^{M}\frac{\hbar^2\nabla_J^2}{M_J} + \frac{1}{2}\sum_{I\neq J}^{M}\frac{e^2Z_IZ_J}{R_{IJ}} - \frac{1}{2}\sum_{j}^{N}\frac{\hbar\nabla_j^2}{m_j} + \frac{1}{2}\sum_{i\neq j}^{N}\frac{e^2}{r_{ij}} - \frac{1}{2}\sum_{i,I}^{N,M}\frac{e^2Z_I}{R_{iI}} \tag{2.2}$$

The many-body Schrodinger equation then can be written as :

$$\left[-\frac{\hbar^2}{2}\sum_i \frac{\nabla^2_{\vec{R}_i}}{M_i} - \frac{\hbar^2}{2}\sum_i \frac{\nabla^2_{\vec{r}_i}}{m_e} - \frac{1}{4\pi\epsilon_0}\sum_{i,j}\frac{e^2 Z_i}{\left|\vec{R}_i - \vec{r}_j\right|} + \frac{1}{8\pi\epsilon_0}\sum_{i\neq j}\frac{e^2}{|\vec{r}_i - \vec{r}_j|} + \frac{1}{8\pi\epsilon_0}\sum_{i\neq j}\frac{e^2 Z_i Z_j}{\left|\vec{R}_i - \vec{R}_j\right|}\right]\Psi(\vec{r},\vec{R}) = E_{tot}\Psi(\vec{r},\vec{R}) \tag{2.3}$$

Equation (2.3) is often referred to as 'the theory of everything' in condensed matter physics [23], given its encompassing nature, covering *almost* everything we need regarding matter's behavior at *equilibrium*. Successfully solving Eq.(2.3) and identifying the eigenfunctions of the system's ground state would open the doors to computing various equilibrium properties. However, the complexity in solving this equation is so daunting that it's borderline impossible so much so that it has been dubbed the " **The many-body problem**", save for the simplest systems e.g. isolated atoms and small molecules. No current computer technology can surpass this limitation. It's possible future advancements might, but it's likely such a machine won't operate on classical principles. Richard P. Feynman captured the essence of this challenge, stating [24]:

> *Nature isn't classical.... and 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.*

To tackle the formidable challenge posed by Eq.(2.3), we adopt a "Divide and Conquer" strategy. This involves utilizing a series of sensible approximations that enable us to concentrate on specific aspects of the problem.

## 2.2 Born-Oppenheimer Approximation

The Schrodinger equation given in Eq.(2.3) can not be solved analytically for multi electron atoms, let alone a solid with Avogadro's number of atoms. One way around this problem is to utilize the inherent difference between the electrons and nuclei masses. The nuclei are three orders of magnitude more massive compared to electrons, thus approximately we can treat the nuclei as static entities. In 1927 Born and Oppenheimer proposed a scheme in which the nuclei can be held fixed in known positions[25], as a result we can neglect their kinetic energy and the potential energy of the repulsion between nuclei can be treated as a constant. Hence, the theoretical picture of the many-body $3M + 3n$ interacting system is reduced to a set of 3n interacting electrons moving in an effective field of the static nuclei. Accordingly Eq.(2.3) is reduced to Eq.(2.4) which is the fundamental equation of *electronic structure theory* :

$$\left[-\frac{\hbar^2}{2}\sum_i \frac{\nabla^2_{\vec{r}_i}}{m_e} + \frac{1}{8\pi\epsilon_0}\sum_{i\neq j}\frac{e^2}{|\vec{r}_i - \vec{r}_j|} + \sum_i V_{ext}\left(\vec{r}_i\right)\right]\psi(\vec{r}) = E\psi(\vec{r}) \tag{2.4}$$

where $V_{ext}$ is the effective field of the static nuclei that is affecting the electrons, it is given by:

$$V_{ext}\left(\vec{r}\right) = -\frac{1}{4\pi\epsilon_0}\sum_{i,j}\frac{e^2 Z_i}{\left|\vec{R}_i - \vec{r}_j\right|} \tag{2.5}$$

Using the Born-Oppenheimer approximation reduces the complexity of the problem, however Eq.(2.4) is still a many-electron problem that is formidable to solve.

## 2.3 Independent Electrons Approximation

The main obstacle in solving the many-electron Schrodinger equation, as denoted in Eq.(2.4) lies in the electron-electron interaction term. The complexity of this term stem from the fact that it requires accounting for the influence of all electrons on one another. Furthermore, the many electron wave function $\psi(\vec{r}_1, ..., \vec{r}_N)$, encompasses the coordinates of all electrons. Given these challenges, direct solution of Eq.(2.4) is infeasible without invoking some sort of approximation to the many-electron wavefunction.

### 2.3.1 Hartree Approximation

In 1928, Hartree put forward the clever **Self-Consistent Field (SCF)** method as an approximate solution for calculating wavefunctions and energies of many electron systems. This method simplifies the complex interactions by assuming electrons move independently, as a result the many-electron wavefunction $\psi_e(\vec{r})$ can be written as a product of one-electron wavefunctions.

$$\psi_H(\vec{r}_1, ...., \vec{r}_N) = \psi_1(\vec{r}_1)\psi_2(\vec{r}_2)....\psi_N(\vec{r}_N) \tag{2.6}$$

Hartree obtained $\psi_H(\vec{r})$ by utilizing the variational principle, which states that

$$\langle\psi_T|\,\hat{H}\,|\psi_T\rangle \geq E_o \tag{2.7}$$

In this framework, $\psi_H(\vec{r})$ minimizes the total energy's expectation value, under the constraint that they are orthonormal. Hartree deduced that the total energy is minimized when the one-electron wavefunctions $\psi_i(\vec{r}_i)$, satisfy the following equation,

known as the Hartree equation:

$$H_H\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r}) \tag{2.8}$$

$$H_H = \frac{-\hbar^2}{2m}\nabla^2 + V_{ext}(\vec{r}) + V_H(\vec{r}) \tag{2.9}$$

where $V_H$ represents the Hartree potential, essentially the "average" potential felt by each electron due to the other electrons. This concept underlies the Hartree method, This concept underlies the Hartree method, commonly referred to as the **mean-field approximation**. The formula for the Hartree potential, $V_H$, is as follows:

$$V_H(\vec{r}) = \frac{e^2}{4\pi\epsilon_o}\int \frac{n(\vec{r'})}{|\vec{r} - \vec{r'}|}d\vec{r'} \tag{2.10}$$

The Hartree approximation simplifies the complexity of the many-electron problem into a more manageable one-electron Schrodinger equation with an effective field ($Veff = V_H + Vext$). However, there is a catch. The catch is that the Hartree potential can only be derived using the electron density $n(\vec{r})$, which is essentially the sum of the probabilities of finding electrons in each occupied state j, given by the equation:

$$n(\vec{r}) = \sum_j |\psi_j(\vec{r})|^2 \tag{2.11}$$

The electron density is described in terms of the wavefunctions, yet these are the very solutions we seek from the Hartree equation. To solve the Hartree equation, you first need to set up the Hamiltonian. This is where the saying " **You need the answer, in order to pose the question**" rings true. To escape this loop, Hartree introduced the Self-Consistent Field (SCF) method. In this method, first you start with a set of of trial wavefunctions $\psi_i(\vec{r_i})$ to create the Hartree potential $V_H$, as per Eq.(2.10). Then, Eq.(2.8) transforms into a single-particle Schrodinger equation, which allows for the calculation of $\psi_i(\vec{r_i})$. After acquiring these new wavefunctions,

$V_H$ is recalculated, and Eq.(2.8) is solved again. This process is repeated until the solutions align closely enough, hence the term "self-consistent" in the method's name.

### 2.3.2 Hartree-Fock Approximation

A key limitation of the Hartree method is that it doesn't take into account the Pauli exclusion principle. The wave function as formulated in Eq.(2.6) does not comply with this fundamental principle of quantum mechanics. Given that electrons are fermions, their *overall* wave function must exhibit antisymmetry when any two electrons are interchanged in terms of their spatial (r) and spin (s) coordinates. This requirement is expressed as:

$$\psi(\vec{r}_1 s_1, ...., \vec{r}_i s_i, ...., \vec{r}_j s_j, ...., \vec{r}_N s_N) = -\psi(\vec{r}_1 s_1, ...., \vec{r}_j s_j, ...., \vec{r}_i s_i, ...., \vec{r}_N s_N) \quad (2.12)$$

To address this issue, a modification to the Hartree wave functions is needed to incorporate the antisymmetry Eq.(2.12). This can be enforced by constructing a Slater determinant, an NxN matrix [26], as the trial wave function. The determinant's property of changing sign upon the interchange of any two columns ensures that the antisymmetry condition is met, thereby aligning the wave function with the Pauli exclusion principle.

$$\psi(\vec{r}_1 s_1, ..., \vec{r}_N s_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(\vec{r}_1 s_1) & \psi_1(\vec{r}_2 s_2) & \cdots & \psi_1(\vec{r}_N s_N) \\ \psi_2(\vec{r}_1 s_1) & \psi_2(\vec{r}_2 s_2) & \cdots & \psi_2(\vec{r}_N s_N) \\ \vdots & \vdots & & \vdots \\ \psi_N(\vec{r}_1 s_1) & \psi_N(\vec{r}_2 s_2) & & \psi_N(\vec{r}_N s_N) \end{vmatrix} \quad (2.13)$$

The extension of the Hartree method to incorporate antisymmetry is known as the Hartree-Fock (HF) method [27]. It essentially follows the Hartree method's procedure but employs an anti-symmetric wavefunction in the variational principle Eq.(2.7).

This anti-symmetric wave function is energetically optimized when it satisfies the Hartree-Fock equation:

$$[\frac{-\hbar^2}{2m}\nabla^2 + V_{ext}(\vec{r}) + V_H(\vec{r})]\psi_i(\vec{r}) - \frac{e^2}{4\pi\epsilon_o}\sum_j \int \frac{1}{|\vec{r}-\vec{r'}|}\psi_j^*(\vec{r'})\psi_i(\vec{r'})\psi_j(\vec{r})\delta_{s_i s_j} d\vec{r'} = \varepsilon_i\psi_i(\vec{r}) \quad (2.14)$$

where $V_H$ is analogous to the Hartree potential. The novel term on the left-hand side, often referred to as the exchange term or the Fock exchange potential, is characterized by the Kronecker-delta $\delta_{s_i,s_j}$, which equals 1 for parallel spins and 0 otherwise. Consequently, the exchange term significantly influences only parallel spins. This term is critical as it embodies the stabilizing interaction that keeps electrons with the same spin further apart, effectively reducing electron-electron Coulomb repulsion. This phenomenon arises directly from the Pauli exclusion principle, which dictates that no two electrons can occupy the same quantum state. Thus, the Hartree-Fock method represents a shift from the classical mean-field approximation to a more quantum-mechanical treatment of electron interactions. However, the inclusion of the exchange term adds complexity to solving HF equations. This complexity primarily stems from the integral nature of the exchange operator, rendering the HF equations more challenging to solve, as noted in [28]. While the Hartree-Fock (HF) method significantly improves upon the Hartree approach by incorporating the Pauli exclusion principle, it still doesn't fully capture the intricate nature of many-electron systems. HF treats electrons as if they're moving independently from each other, only experiencing Coulombic repulsion that's averaged over the positions of all electrons. This approximation overlooks the correlation in electron positions due to their Coulomb interaction. This oversight is particularly critical for electron pairs with anti-parallel spins, as the exchange term in HF doesn't help in keeping them apart [29].

The idea behind the independent electron approximation is to simplify the many-body problem by estimating the complex, many-body wavefunction. However, this approach becomes extremely challenging, both conceptually and numerically, especially in complex systems like solids that contain a vast number of atoms.

Given these limitations, there's a need to explore alternative methods that don't rely on wavefunctions as the primary variable. These alternative approaches aim to provide a more manageable yet accurate description of electron behavior in many-electron systems, particularly in complex materials.

## 2.4 Density Functional Theory

Density Functional Theory (DFT) stands out as one of the most favored and precise approaches for first-principles calculations concerning the electronic structure of atoms, molecules, crystals, and surfaces. Its prominence is further highlighted by its academic impact: twelve of the top-100 most cited papers in the physical sciences are related to DFT, including two in the top 10 [30]. This level of citation not only underscores DFT's importance but also its status as a successful alternative to wavefunction-based methods like the Hartree-Fock approximation [31].

In 1998, Walter Kohn was awarded the Nobel Prize for his pivotal contributions to the development of Density Functional Theory (DFT) [32]. The fundamental insight of DFT is that the ground state energy, E, is a functional of the electron density $\mathrm{n}(\vec{r})$, represented as:

$$E = F[n(\vec{r})] \tag{2.15}$$

This concept is particularly noteworthy because $\mathrm{n}(\vec{r})$ depends only on three spatial coordinates, in stark contrast to the multi-electron wavefunction, which involves a significantly more complex dependency on 3N variables (with N being the number of

electrons). Therefore, DFT simplifies the problem by abandoning the many-electron wavefunction in favor of focusing on the electron density. The central variable in DFT calculations, then, is this electron density n($\vec{r}$), which for single particles can be defined as:

$$n(\vec{r}) = \sum_{j} |\psi_j(\vec{r})|^2 \tag{2.16}$$

This shift to working with the electron density, a function of just three variables, instead of a highly complex many-electron wavefunction that involves 3N variables, is a major advantage of DFT. It significantly simplifies the analysis of a wide range of systems. While the theory itself is conceptually sound and theoretically precise, practical applications require approximations. These approximations are necessary because, despite the conceptual elegance of DFT, the exact functional forms required to fully implement the theory are not known. This section will further explore the formulation of DFT and how it addresses these challenges.

### 2.4.1 Thomas-Fermi Model

The Thomas-Fermi (TF) model, historically regarded as the conceptual root for Density Functional Theory (DFT), was first proposed by Thomas and Fermi [33, 34]. They pioneered the idea of utilizing electron density to describe electron distribution within a system. In their model, they considered electrons interacting within an external potential $v_{ext}$, proposing an over-simplified relationship between this external potential and the electron density. Their model expressed the relationship between the interacting electrons in the external potential and the electron density n($\vec{r}$) as:

$$n(\vec{r}) = \gamma[\mu - v_{eff}(\vec{r})]^{\frac{3}{2}} \tag{2.17}$$

where $\gamma = \frac{1}{3\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}$, $\mu$ is the chemical potential and $v_{eff}$ is the external potential plus the classical electrostatic potential generated by the electron density:

$$v_{eff} \equiv v_{ext}(\vec{r}) + \frac{e^2}{4\pi\epsilon_o}\int \frac{n(\vec{r}')}{|\vec{r} - \vec{r}'|}d\vec{r}' \tag{2.18}$$

Although the TF model marks a significant milestone on the path to DFT, its practical applications are quite restricted. It fails to account for many crucial phenomena, primarily because it is a basic approximation that overlooks exchange and correlation effects. Walter Kohn, in his Nobel lecture [32], reflected on the Thomas-Fermi model, noting its limitations and the groundwork it laid for the development of more advanced theories like DFT.

> *... This raised a general question in my mind: Is a complete, exact description of the electronic structure in terms of n(r) possible in principle. A key question was whether the density n(r) completely characterized the the system. This was true in TF theory.*
>
> *... This suggested the hypothesis that a knowledge of the ground state density of n(r) for any electron system, with or without interactions, uniquely determines the system. This hypothesis became the starting point of modern DFT.*

### 2.4.2 The Hohenberg - Kohn Theorem

Hohenberg and Kohn, in 1964 [35], established two fundamental theorems that are the cornerstone of Density Functional Theory (DFT). Their aim was to develop DFT as an **exact** framework for understanding many-body systems [36]. These theorems articulate that:

**Theorem I**

*For any system of interacting particles in an external potential $V_{ext}$, the ground state particle density $n_o(\boldsymbol{r})$ uniquely determines this external potential.*

This theorem underlines a crucial one-to-one correspondence between the ground state density and the external potential:

$$n(\vec{r}) \Leftrightarrow V_{ext}(\vec{r})$$

A key consequence of this is that the Hamiltonian of a many-electron interacting system Eq.(2.19) is also uniquely defined by the ground-state density.

$$\hat{H} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ext} \tag{2.19}$$

Here, the first two terms account for the kinetic energy and electron-electron interactions, respectively, and are considered universal for all many-electron systems. The last term, involving the external potential $V_{ext}$ encapsulate all the system-specific details. From this, we infer that the total energy (E) of any many-electron system uniquely depends on the electron density $n(\vec{r})$ , formulated as $E = E[n(\vec{r})]$. Breaking down the total energy functional, we have:

$$E[n(\vec{r})] = \langle \psi[n] | \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ext} \, | \psi[n] \rangle \tag{2.20}$$

Further, it can be expressed as:

$$E[n] = T_e[n] + V_{ee}[n] + \int V_{ext}(\vec{r}) n(\vec{r}) d\vec{r} \tag{2.21}$$

The first two universal terms are often combined into the Hohenberg-Kohn universal functional $F_{HK}[n] = T_e[n] + V_{ee}[n]$. The energy functional can be expressed as:

$$E[n] = F_{HK}[n] + \int V_{ext}(\vec{r})n(\vec{r})d\vec{r} \tag{2.22}$$

Thus, understanding and determining the ground-state density $n(\vec{r})$ becomes central to characterizing the entire system under study in DFT. Unfortunately, a significant challenge in DFT is the unknown true form of the universal functional $F_{HK}[n]$ for many-electron systems, making approximations crucial for practical applications.

**Theorem II**

*The ground state density of the interacting system associated with the unique $V_{ext}$ minimizes the total energy of that system.*

Diving deeper, Theorem II reveals the true nature of the energy functional in DFT as inherently variational. This essentially means that, provided we knew the exact form of this functional, it could be harnessed through the variational principle to deduce the electron density that minimizes it, thereby pinpointing the system's true ground state density. This principle, often referred to as the **"Hohenberg-Kohn variational principle"**, can be expressed as:

$$E_o \leq E[n] = T_e[n] + V_{ee}[n] + V_{ext}[n] \tag{2.23}$$

$$\left.\frac{\delta E[n]}{\delta n}\right|_{n_o} = 0 \tag{2.24}$$

While the Hohenberg-Kohn (HK) theorems brilliantly set the stage by establishing a unique link between $n(\vec{r})$ and $V_{ext}$ and framing the ground state energy as a functional of $n(\vec{r})$, they stop short of providing a blueprint for constructing this

elusive functional. This is where the ingenuity of the Kohn-Sham approach becomes pivotal, bridging the gap left by the HK theorems and offering a practical pathway to apply DFT in real-world scenarios.

### 2.4.3 The Kohn- Sham Approach

In addressing the practical challenges posed by the Hohenberg and Kohn theorems, Kohn and Sham [37], introduced a revolutionary method that fundamentally changed how we approach Density Functional Theory (DFT). Their ingenious contribution lay in simplifying the daunting many-electron problem. Instead of directly dealing with a system of interacting electrons, they proposed a scenario where these electrons are considered non-interacting but moving in an effective potential, as illustrated in Fig.2.2. This effective potential, $V_{KS}$, is designed to mimic the behavior of the interacting electrons so closely that it yields the same ground state electron density.

In addition to this, the wavefunctions $\phi(\vec{r})$, referred to as KS wavefunctions in this non-interacting model, need to mimic the ground state density of the interacting system. These wavefunctions are tailored to yield the same electron density as the true interacting system, expressed as:

$$n(\vec{r}) = \sum_i |\phi_i(\vec{r})|^2 \tag{2.25}$$

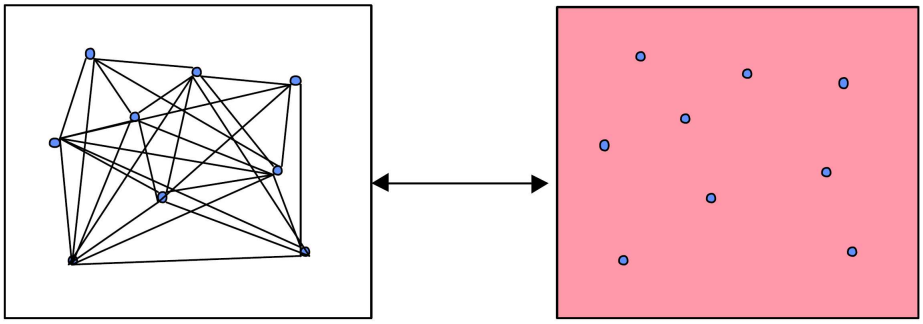


**Figure 2.2:** An illustration of the relationship between the "real" many-electron system, LHS, and the non-interacting system of Kohn Sham density functional theory, RHS [38].

Consequently, the energy functional E[n] for the interacting system can be decomposed to a non-interacting part, similar to the Hartree approximation and an additional term that encapsulates the complexities of the many-body problem. The functional E[n] is given by:

$$E[n] = T_{KS}[n] + \frac{e^2}{8\pi\varepsilon_o}\int\int \frac{n(\vec{r})n(\vec{r^*})}{|\vec{r}-\vec{r^*}|}d\vec{r}d\vec{r^*} + \int V_{ext}(\vec{r})n(\vec{r})d\vec{r} + E_{XC}[n] \quad (2.26)$$

where $T_{KS}$ is the kinetic energy of the fictitious system of non-interacting electrons. It is formulated as:

$$T_{KS} = \frac{-\hbar^2}{2m}\sum_i \langle\phi_i|\nabla^2|\phi_i\rangle \quad (2.27)$$

Each $\psi_i$ represents a Kohn-Sham orbital, the summation runs over all the orbitals emphasizing the independent particle approximation. The second term represents the Hartree potential energy $U_H$, while the third term represents the external potential functional. The last term known as the exchange-correlation functional $E_{XC}$ is a the crucial component that encapsulates the difference between the true quantum mechanical interactions in the system and the approximations made by the first

three terms.

$$\begin{aligned} E_{XC}[n] &= F_{HK}[n] - T_{KS}[n] - U_H[n] \\ &= T_e[n] + V_{ee}[n] - T_{KS}[n] - U_H[n] \end{aligned} \tag{2.28}$$

This term accounts for the exchange and correlation effects in both the kinetic and potential energy.

Applying the Hohenberg-Kohn variational principle Eq.(2.24), leads to the Kohn-Sham equations, which governs the KS wavefunctions. These are single-particle Schrodinger-like equations:

$$\left[\frac{-\hbar^2}{2m}\nabla_i^2 + \hat{V}_{ext} + \hat{V}_H + \hat{V}_{XC}\right]\phi_i(\vec{r}) = \varepsilon_i\phi_i(\vec{r}) \tag{2.29}$$

Where, $\hat{V}_{ext}$ is the external potential Eq.(2.5), $\hat{V}_H$ is the Hartree mean-field potential Eq.(2.10), and $\hat{V}_{XC}$ is the exchange and correlation potential, defined as:

$$V_{XC} = \frac{\delta E_{XC}[n(\vec{r})]}{\delta n(\vec{r})} \tag{2.30}$$

The Kohn-Sham equation Eq.(2.29) is the cornerstone of DFT, enabling the calculation of electron density in a self-consistent manner, as both $\hat{V}_H$ and $\hat{V}_{XC}$ depend on $n(\vec{r})$.

In principle the Kohn-Sham method stands as an exact theoretical framework, remarkable in its lack of inherent approximations. It enables the accurate determination of the electron density for an interacting system through the solutions of a self-consistent set of single-particle, Schrodinger-like equations, which is a huge feat. A major drawback however, is that the exact form of the exchange-correlation energy functional $E_{XC}$ remains unknown. If we could obtain the precise form of $E_{XC}$, then solving the Kohn-Sham equations would yield the ground state density which is identical to that of the many-electron problem, unlocking a deeper understanding of the system's quantum state. In the quest to overcome this obstacle, the universality of $E_{XC}$ offers

a silver lining. Obtaining an effective approximation for $E_{XC}$ would enable us to solve the Kohn-Sham equations for virtually any many-electron system, revealing both the ground state density and total energy. This potential was eloquently highlighted in the Nobel Prize lecture by W. Kohn [32],

> *The KS theory may be regarded as the formal exactification of Hartree theory. With the exact $E_{XC}$ and $V_{xc}$ all many-body effects are in principle included. Clearly this directs attention to the functional $E_{XC}[n(\vec{r})]$. The practical usefulness of ground-state DFT depends entirely on whether approximations for the functional $E_{XC}[n(\vec{r})]$ could be found which are at the same time sufficiently simple and sufficiently accurate.*

Effectively using the Kohn-Sham equations to precisely solve the many-electron problem hinges on understanding the exchange-correlation functional $E_{XC}[n(\vec{r})]$. As it stands, our incomplete understanding of $E_{XC}[n(\vec{r})]$ limits the ability to exploit the full potential of the Kohn-Sham approach for **exact** solutions. This has led to the development of various approximations for $E_{XC}[n(\vec{r})]$, making it a key area of ongoing research.

### 2.4.4 Exchange-Correlation Functional

As previously stated, the exact nature of $E_{XC}[n(\vec{r})]$functional remains elusive. The choice of approximation for this unknown functional significantly dictates the accuracy of DFT calculations. Therefore, the development of refined approximations for $E_{XC}[n(\vec{r})]$ remains as one of the greatest challenges in modern DFT and continues to be an active point of research. Among the several approximations put forth, the

Local Density Approximation (LDA) and the Generalized Gradient Approximation (GGA) are particularly noteworthy.

In the LDA framework, $E_{XC}^{LDA}[n(\vec{r})]$ mirrors the characteristics of a homogeneous electron gas:

$$E_{XC}^{LDA}[n(\vec{r})] = \int n(\vec{r})\varepsilon_{XC}(n(\vec{r}))d\vec{r} \tag{2.31}$$

Here, $\varepsilon_{XC}$ denotes the exchange-correlation energy density for a uniform electron gas at density n($\vec{r}$) [37]. LDA operates under the assumption of a slowly varying electron density, akin to a locally uniform electron gas. Despite delivering generally reasonable results, LDA tends to overestimate lattice constants [39]. The Generalized Gradient Approximation (GGA), on the other hand, aims to enhance LDA by incorporating a density gradient [40]. Which means unlike LDA, GGA addresses non-uniform charge densities.

$$E_{XC}^{GGA}[n(\vec{r})] = \int n(\vec{r})\varepsilon_{XC}(n(\vec{r}), \nabla n(\vec{r}))d\vec{r} \tag{2.32}$$

One widely used formulation within GGA is the PBE-GGA by Perdew et al. [41]. GGA generally outperforms LDA, especially in predicting structural and magnetic properties, as the results align closer to experimental values [39].

### 2.4.5 Self-Consistent Calculations

In our exploration of Density functional theory (DFT) and the Kohn-Sham approach, we have arrived at a set of crucial equations: Eq.(2.29) and Eq.(2.25). These equations are intricately linked, as both $V_H$ and $V_{XC}$ are dependent on the electron density $n(\vec{r})$. This interdependence necessitates a self-consistent solution approach. To find the accurate electron density n($\vec{r}$), we must iteratively solve Equations (2.29) and (2.25), refining our results in each iteration until self-consistency is achieved. The process of solving the Kohn-Sham equations is more of a **" trial and error "** method. Fig.2.3 offers a visual representation of the iterative cycle, illustrating the

step-by-step process involved in reaching self-consistency.

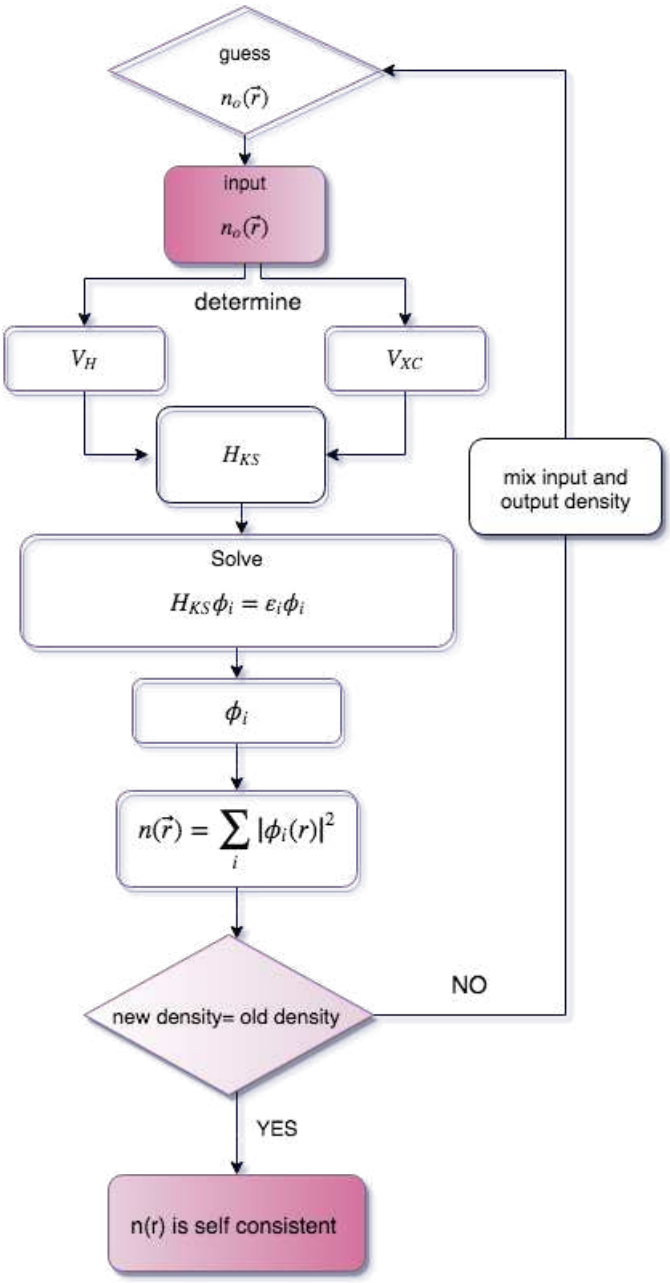



**Figure 2.3:** Flow chart for finding self-consistent solutions of KS equations.

It is worth noting that DFT calculations scale as $O(N^3)$ with the system size N typically indicated by the number electrons. This is due to the matrix diagonalization of the Kohn-Sham Hamiltonian which is required in a straightforward implementation [42, 43].

### 2.4.6 Limitations of DFT

One major limitation of Kohn-Sham density functional theory is its inability to accurately predict the band gap, a challenge so significant within the computational solid state community that it has been dubbed '**The band gap problem**'. The precise calculation of the band gap is not just a theoretical challenge; it holds practical importance. The band gap is a critical factor in the performance and the efficiency of various technological applications, particularly in opto-electronics and photovoltaics [44].

In theory, DFT can determine the fundamental gap by calculating the difference in total ground state energies after adding and removing electrons:

$$E_{gap} = [E(N+1) - E(N)] - [E(N) - E(N-1)] \tag{2.33}$$

where the particle number of the system is indicated in parenthesis. Essentially, the band gap is the difference between electron affinity $[E(N+1) - E(N)]$ and the ionization energy $[E(N) - E(N-1)]$. A key issue arises when one tries to ascribe physical meaning to the Kohn-Sham eigenvalues. It is possible to do that for the highest occupied eigenvalue $\varepsilon_N$, as proven by Janak [45], Where the ionization energy of the system is equal to the negative of the highest occupied eigenvalue $[I(N) = -\varepsilon_N]$. The Kohn-Sham band gap calculated using the the same N-electron Kohn-Sham Hamiltonian is given by the difference between the lowest unoccupied eigenvalue and the highest occupied eigenvalue:

$$\varepsilon_g^{KS} = \varepsilon_N(N+1) - \varepsilon_N(N) \tag{2.34}$$

Expressing the band gap in terms of Kohn-Sham eigenvalues:

$$E_g = \varepsilon_{N+1}(N+1) - \varepsilon_N(N) \tag{2.35}$$

It becomes apparent that the two quantities in Eq.(2.35) and Eq.(2.34) are not the same, differing by the so called derivative discontinuity [$\Delta = \varepsilon_{N+1}(N+1) - \varepsilon_{N+1}(N)$] [46]. This discrepancy arises because E(N) changes slope at integer particle number.

$$E_g = \varepsilon_g^{KS} + \Delta \tag{2.36}$$

Thus, to accurately utilize the Kohn-Sham eigenvalues in band gap calculations, both terms in Eq.(2.36) must be present. The band gap problem is not an inherent flaw of the theory but rather it is a case of misuse of the Kohn-Sham equations, expecting them to calculate quantities beyond their theoretical reach [47]. To overcome this limitation, we must look beyond the conventional KS-DFT framework.

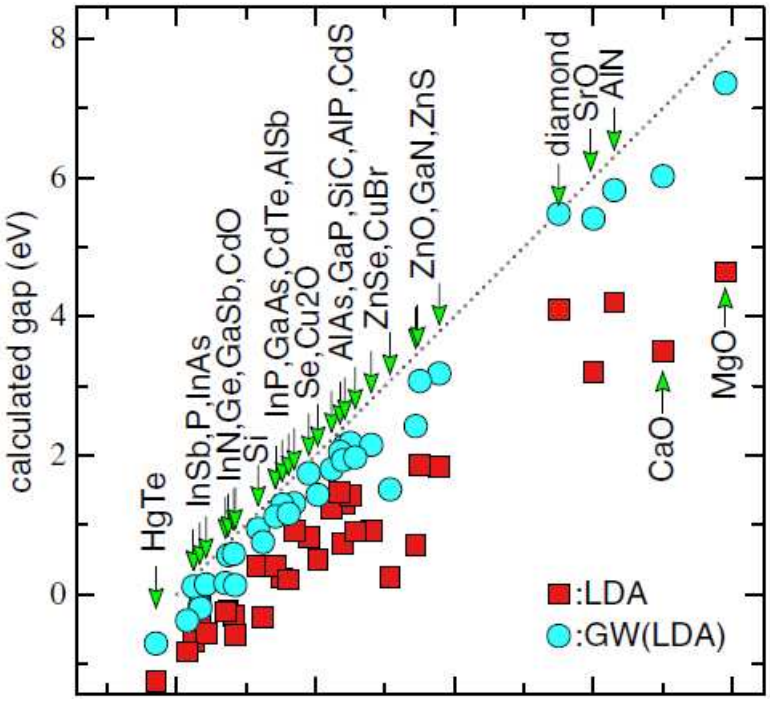



**Figure 2.4:** DFT underestimate the band gap [4]

# Chapter Three

# The Realm Of Excited States

DFT, by design a ground-state theory, ideally, is not suited to describe excitation in many-electron systems. However, it can serve as a stepping stone to cross into the realm of excited states. When crossing one needs to distinguish between two classes of excitation. First, the charged ones, where the total number of electrons differs before and after the perturbation. The second class is neutral excitation, where the total number of electrons doesn't change upon excitation.[1]

## 3.1 Many-body perturbation theory

In principle one could start from an independent particle picture and consider the Coulomb interaction as a perturbation. However, one would be faced with two major issues; the Coulomb interaction is not small by any means, and it has many effects that would be absent if the term was neglected. The other major issue being that Coulomb interaction is a two-body effect, which means the problem scales badly with the number of electrons making it intractable for large systems [47]. This means that we need to reformulate the problem in order to obtain a well defined theory.

Taking inspiration from the success that Kohn-Sham approach had in describing the ground state properties, where they mapped the interacting electron system into a fictitious one of non-interacting particles in an effective potential $V_{ext}$.

## 3.2 Charged excitation: The quasiparticle picture

Fig.3.1 depicts the process of photo-induced charged excitation. In the independent particle picture, upon excitation a hole is formed in the electron sea, the formation of this hole has no effects on its surrounding as previously outlined in Sec.1.2.1.

[1]Note that all equations in the following chapter make use of Hartree atomic units unless stated otherwise ($m_e = \hbar = e = 4\pi\epsilon_0 = 1$)

However, we know that is not the case. In reality, the hole gets dynamically screened by a cloud of the surrounding electrons. The hole plus its screening cloud is known as a quasi-particle. Same concept applies to adding an extra electron to the system.

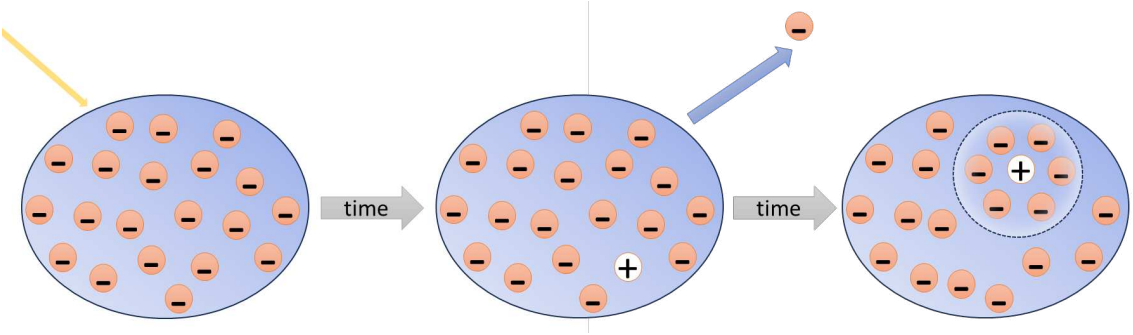



**Figure 3.1:** An illustration of screening and quasiparticle formation: starting from photon-induced excitation up to the creation of a quasiparticle within the electron sea. The screening of the positive charge is based on the simple idea that charges in the system rearrange themselves in such a way as to minimize interactions.

To describe the quasiparticle picture of charged excitation within many-body perturbation theory, we begin by discussing its building blocks, namely Green's functions.

## 3.3 One-particle Green's function

The key ingredient in many-body perturbation theory is the one-particle Green's function. It is defined as follows [48]:

$$G(r,t;r',t') = -i \left\langle \Psi_0 \right| T \left[ \hat{\psi}_H(r,t) \hat{\psi}_H^\dagger(r',t') \right] \left| \Psi_0 \right\rangle \tag{3.1}$$

where $|\Psi_0\rangle$ is the ground state of the interacting system, which is assumed to be normalized. The T product is the time ordering operator, which ensures that the operators are ordered chronologically from right to left and a minus sign should be included when fermion operators are interchanged to obtain the right order.

Explicitly, it can be written as follows:

$$\begin{cases} G(r,t;r',t') = -i \left\langle \Psi_0 \right| \left[ \hat{\psi}_H(r,t) \hat{\psi}_H^\dagger(r',t') \right] \left| \Psi_0 \right\rangle & t > t' \\ G(r,t;r',t') = i \left\langle \Psi_0 \right| \left[ \hat{\psi}_H^\dagger(r',t') \hat{\psi}_H(r,t) \right] \left| \Psi_0 \right\rangle & t < t' \end{cases}$$

$\hat{\psi}_H(r,t)$, $\hat{\psi}_H^\dagger(r',t')$ are the Heisenberg field operators. To understand the physical interpretation of Green's function and how it can be seen as particle propagator, we expand the time dependence in Heisenberg field operators:

$$\hat{\psi}_H = e^{i\hat{H}t} \psi e^{-i\hat{H}t} \tag{3.2}$$

The Green's function with explicit time dependence can be written as:

$$\begin{aligned} G(rt;r't') = &-i\Theta(t-t') \left\langle \Psi_0 \right| \psi(r) e^{-i(H-E_0)(t-t')} \psi^\dagger(r') \left| \Psi_0 \right\rangle \\ &+ i\Theta(t'-t) \left\langle \Psi_0 \right| \psi^\dagger(r') e^{i(H-E_0)(t-t')} \psi(r) \left| \Psi_0 \right\rangle \end{aligned} \tag{3.3}$$

where $\Theta$ is the heavy side function, for $(t > t')$ the field operator acting on the ground state $\psi^\dagger(r') \left| \Psi_0 \right\rangle$ will result on a creation of an electron at position r at time $t'$, that will propagate in time according to $e^{-i(H-E_0)(t-t')/\hbar}$ projecting this onto the state $\left\langle \Psi_0 \right| \psi(r)$ which is the ground state with an added electron at time t at position r.

The first part of Green's function in Eq.3.3 characterizes the propagation of a state containing one extra electron. For $(t' > t)$, in a similar manner it, Green's function represent the propagation of a state with an extra hole.

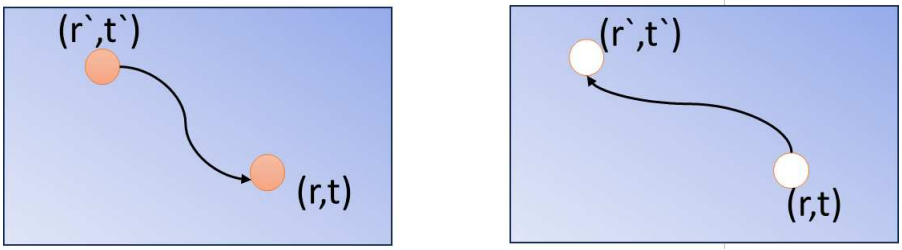



**Figure 3.2:** A sketch representing the propagation of an electron (left) and a hole (right) added to the ground state of the many body system.

The one-body Green's function provides insights into various key properties, such as the expectation values of one-body operators and the spectral function, which is related to photoemission spectra [49]. Additionally, one important property that can be formulated in terms of the One-body Green's function is the total energy, which is given by the Galitskii–Migdal expression [50]

$$E_0 = \frac{1}{2}\int dr \lim_{r'\to r}\lim_{t'\to t^+}\left[\frac{\partial}{\partial t} - ih_0(r)\right] G(r,r',t,t') \tag{3.4}$$

where $h_0(r)$is the non interacting part of the Hamiltonian.

The one-body Green's function can be written in a spectral representation also known as the Lehmann representation [51]. It can be obtained by inserting a complete set of $(N \pm 1)$ states in between the field operators in Eq.3.3 and performing Fourier transformation. Assuming a time-translational invariance, the Fourier transform can be taken with respect to (t-t') which gives [52]:

$$G(r,r',\omega) = \sum_s \frac{\left\langle \Psi_0^N \right| \psi(r) \left| \Psi_S^{N+1} \right\rangle \left\langle \Psi_S^{N+1} \right| \psi^\dagger(r') \left| \Psi_0^N \right\rangle}{\omega - (E_S^{N+1} - E_0^N) + i\eta} \tag{3.5}$$

$$+ \frac{\left\langle \Psi_0^N \right| \psi^\dagger(r) \left| \Psi_S^{N-1} \right\rangle \left\langle \Psi_S^{N-1} \right| \psi(r') \left| \Psi_0^N \right\rangle}{\omega + (E_S^{N-1} - E_0^N) - i\eta} \tag{3.6}$$

where $\eta$ is a positive infinitesimal. $\left|\Psi_s^{N\pm}\right\rangle$ and $E_s^{N\pm}$ are the eigenstates and energy of the sth eigenstate of the $N \pm 1$ electrons system. An important feature of which makes

working with Green's functions even more appealing is the fact the singularities of G in the Lehmann representation gives direct access to the electron's removal and addition energies as measured in photoemsssion experiments Sec.1.2.1.

$$E_{\text{removal}} = E_0^{(N)} - E_s^{(N-1)}$$
$$E_{\text{add}} = E_s^{(N+1)} - E_0^{(N)}$$

The Lehmann representation can also be used to write the Spectral function which is proportional to the imaginary part of the Green's function.

$$A(r,r',\omega) = \sum_s \left\langle \Psi_0^N \right| \psi(r) \left| \Psi_S^{N+1} \right\rangle \left\langle \Psi_S^{N+1} \right| \psi^\dagger(r') \left| \Psi_0^N \right\rangle \delta(\omega - (E_S^{N+1} - E_0^N)) + \sum_s \left\langle \Psi_0^N \right| \psi^\dagger(r) \left| \Psi_S^{N-1} \right\rangle \left\langle \Psi_S^{N-1} \right| \psi(r') \left| \Psi_0^N \right\rangle \delta(\omega + (E_S^{N-1} - E_0^N)) \quad (3.7)$$

The spectral function is a crucial tool as it provides direct information about the single-particle excitation of the system and can be measured experimentally in photoemission or inverse photoemission.

As mentioned before the one-body Green's function $G(rt, r't')$ describes the propagation of a single particle from one space-time point $(r', t')$ to another $(r, t)$. The equation of motion for the Green's function is derived from the Heisenberg picture

$$\partial_t \left[ \Theta(t-t') \left\langle \Psi_0 \right| \hat{\psi}_H(r,t) \hat{\psi}_H^\dagger(r',t') \left| \Psi_0 \right\rangle \right] = \delta(t-t') \left\langle \Psi_0 \right| \hat{\psi}_H(r,t) \hat{\psi}_H^\dagger(r',t') \left| \Psi_0 \right\rangle \quad (3.8)$$
$$+\Theta(t-t') \left\langle \Psi_0 \right| \hat{\partial}_t [\psi_H(r,t) \hat{\psi}_H^\dagger(r',t')] \left| \Psi_0 \right\rangle \quad (3.9)$$

which gives,

$$\left(i\frac{\partial}{\partial t} - h_0(r)\right) G(rt, r't') = \delta(t-t')\delta(r-r') - i\int dr'' v(r-r'') \langle \Psi_0 | T\left[\hat{\psi}^\dagger(r''t^{++})\hat{\psi}(r''t^{+})\hat{\psi}(rt)\hat{\psi}^\dagger(r't')\right] |\Psi_0\rangle \quad (3.10)$$

The infinitesimal increments $++$ and $+$ are used to break the time degeneracy and order the operators correctly. $h_0(r)$ represents the one-body Hamiltonian, including the kinetic energy and the external potential. The integral term includes the effects of two-body interactions through Coulomb interaction $v(r-r'')$. One recognize the two-body Green's function which is given by the expectation value of the time-ordered product of field operators.

$$i^2 G_2(1,3;2,4) = \langle \Psi_0 | T[\hat{\psi}(1)\hat{\psi}(3)\hat{\psi}^\dagger(4)\hat{\psi}^\dagger(2)] | \Psi_0 \rangle \quad (3.11)$$

for brevity and clarity sake we use the shorthand indices 1, 2, 3,4, where $(r,t) \Rightarrow (1)$ and so on. The equation of motion for the one-body Green's function is now given in terms of a two-body Green's function

$$\left(i\frac{\partial}{\partial t_1} - h_0(r_1)\right) G(1,2) + i\int d3 v(1,3) G_2(1,3^+;2,3^{++}) = \delta(1,2) \quad (3.12)$$

with $v(1,3) = v(r_1 - r_3)\delta(t_1 - t_3)$ being a two-body instantaneous interaction. We stumbled upon the Green's function *hierarchy*, known as the Martin–Schwinger hierarchy [53], which states that the equation of motion for a Green's function of order s, $G_s$ leads to the appearance of the Green's function of order $s+1$.

The equation of motion for Green's function is given by

$$\left[i\frac{\partial}{\partial t} - H_0(\mathbf{r},t)\right] G(\mathbf{r},t,\mathbf{r}',t') - \int d\mathbf{r}''dt''\Sigma(\mathbf{r},t,\mathbf{r}'',t'')G(\mathbf{r}'',t'',\mathbf{r}',t') = \delta(\mathbf{r}-\mathbf{r}',t-t') \tag{3.13}$$

we adopt the shorthand notation $(r,t) \Rightarrow (1)$.

$$\left[i\frac{\partial}{\partial t} - H_0(1)\right] G(1,2) - \int d3\Sigma(1,3)G(3,2) = \delta(1,2) \tag{3.14}$$

The Hamiltonian $H_0$ includes the following terms:

$$H_0 = -\frac{1}{2}\nabla_1^2 + V_{ext}(1) + V_H(1)$$

which are in order the kinetic energy of the particles, the static external potential and the Hartree potential.

Taking the self energy $\Sigma = 0$, Eq.3.14 reduces to the equation of motion for the non-interacting Green's function $G_0$

$$\left[i\frac{\partial}{\partial t} - H_0(1)\right] G_0(1,2) = \delta(1,2) \tag{3.15}$$

Starting from the equation of motion of the Green's function Eq.3.14, We can rearrange it to isolate the Green's function $G(1,2)$ by using the equation of motion of Green's function for the non-interacting system Eq.3.15. We get the integral Dyson [54] equation for the Green function:

$$G(1,2) = G_0(1,2) + \int d(34) G_0(1,3)\Sigma(3,4)G(4,2) \tag{3.16}$$

Eq.3.16 describes how the Green's function of the interacting system $G$ is related to the Green's function of the non-interacting system $G_0$. The integral term involving the self-energy $\Sigma$ and $G$ accounts for the perturbation of Green's function $G_0$ due to all possible exchange-correlation interactions that a particle can experience while propagating from point 2 to point 1, so in a sense $\Sigma$ is kind of an effective potential. For a more succinct expression, Eq.3.16 can be expressed as follows:

$$G = G_0 + G_0 \Sigma G \tag{3.17}$$

### 3.3.1 Hedin Equations

Hedin's equations [55] are a set of five integro-differential equations that connect the one-body Green's function G, the vertex function $\Gamma$, the polarizability function P which is the response of the system to a perturbing potential, the dynamically screened Coulomb interaction W, the bare one v and the self-energy $\Sigma$. In principle, this is the full recipe to calculate the self-energy operator $\Sigma$

$$G(1,2) = G_0(1,2) + \int d(34) G_0(1,3)\Sigma(3,4)G(4,2) \tag{3.18}$$

$$\Gamma(1,2,3) = \delta(1-2)\delta(2-3) + \int d(4567) \frac{\delta\Sigma(1,2)}{\delta G(4,5)} G(4,6)G(7,5)\Gamma(6,7,3) \tag{3.19}$$

$$P(1,2) = -i \int d(34) G(1,3)\Gamma(324)G(4,1^+) \tag{3.20}$$

$$W(1,2) = v(1,2) + \int d(34) v(1,3)P(3,4)W(4,2) \tag{3.21}$$

$$\Sigma(1,2) = i \int d(34) G(1,3^+)W(1,4)\Gamma(3,2,4) \tag{3.22}$$

The index (1) is a shorthand notation for position,time and spin coordinates: $(1) = (r_1, t_1, \sigma_1)$ similarly for the other indices, while $(1^+)$ represents $(r_1, t_1 + \delta)$ with $\delta$ being a positive infinitesimal. The Hedin's equations are solved self-consistently. Fig.3.3 shows the Hedin's pentagon, which illustrates the flowchart for the iterative solution.

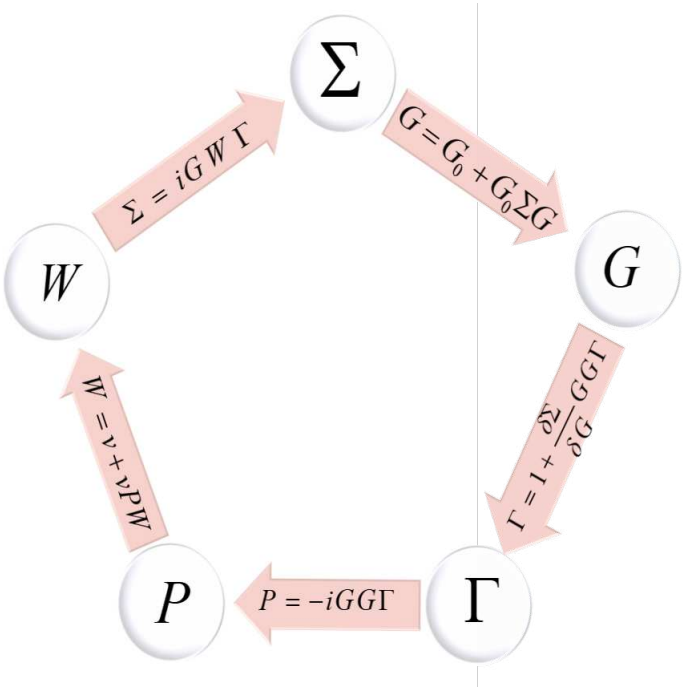



**Figure 3.3:** Hedin's pentagon [55]

### 3.3.2 From Hedin's equations to GW approximation

The set of Hedin's equations, in their original form, are exact but challenging to solve. The standard approach is to use the GW approximation [56, 55, 52], which proved to be quite successful [57, 58, 59, 60]. In this approach, the self energy is approximated to the lowest order with respect to the screened Coulomb interaction W, hence the name "**GW**" approximation. This is done by taking the

$$\Sigma = iGW \tag{3.23}$$

This is done by only considering the trivial delta function contribution in the vertex function Γ, Eq.3.19 and ignoring the second term. As a result, Hedin's equations (Eq.3.18-Eq.3.28) Within the GW approximation reduce to:

$$G(1,2) = G_0(1,2) + \int d(34) G_0(1,3)\Sigma(3,4)G(4,2) \tag{3.24}$$

$$\Gamma(1,2,3) = \delta(1-2)\delta(2-3) \tag{3.25}$$

$$P(1,2) = -iG(1,2)G(21) \tag{3.26}$$

$$W(1,2) = v(1,2) + \int d(34) v(1,3)P(3,4)W(4,2) \tag{3.27}$$

$$\Sigma(1,2) = iG(1,2^+)W(1,2) \tag{3.28}$$

These equations are used to obtain the self-energy operator [61].

## 3.4 GW in Practice

In this section we give an overview of the approach followed in [62, 63]. The irreducible dynamical polarizability within the random phase approximation for the spin-unpolarized case, is written as

$$P^o(\boldsymbol{r}, \boldsymbol{r}'; i\omega) = 4\,\mathrm{Re}\sum_{cv} \frac{\psi_c(\boldsymbol{r})\psi_v(\boldsymbol{r})'\psi_v(\boldsymbol{r})\psi_c(\boldsymbol{r}')}{i\omega - (\varepsilon_c - \varepsilon_v)} \tag{3.29}$$

which can be expanded using the optimal polarizability basis $|\Phi_\mu\rangle$ described in [62].

$$P^o(\mathbf{r}, \mathbf{r}'; i\omega) = \sum_{\mu\nu} \Phi_\mu(\mathbf{r}) P^o_{\mu\nu}(i\omega)\Phi_\nu(\mathbf{r}') \tag{3.30}$$

where matrix elements of the polarizability operator $P^o_{\mu\nu}$ in the basis set $\Phi_\mu$ are given by:

$$P^o_{\mu\nu}(i\omega) = -4\,\mathrm{Re}\sum_{vc}\int d\mathbf{r}\mathbf{r}'\Phi_\mu \frac{\psi_v(\mathbf{r})\psi_c(\mathbf{r})\psi_v(\mathbf{r}')\psi_c(\mathbf{r}')}{(\varepsilon_c - \varepsilon_v) + i\omega}\Phi_\nu(\mathbf{r}') \tag{3.31}$$

Motivated by Sternheimer approach [64], Umari et.al [63] introduced the projector operator over the conduction manifold $P_c$, to re-write the polarizability matrix elements in Eq.3.37 without any explicit reference to the empty states. The projector operator $P_c$ can be written as:

$$P_c(\boldsymbol{r},\boldsymbol{r}') = \sum_c \psi_c(\boldsymbol{r})\psi_c(\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}') - \sum_v \psi_v(\boldsymbol{r})\psi_v(\boldsymbol{r}') \tag{3.32}$$

with the notation:

$$\langle \mathbf{r}|\ \psi_i\Phi_v\rangle = \psi_i(\mathbf{r})\Phi_v(\mathbf{r}) \tag{3.33}$$

the matrix elements read:

$$P^o_{\mu\nu}(i\omega) = -4\,\mathrm{Re}\left[\sum_v \left\langle \psi_v\Phi_\mu \left| \hat{P}_c \frac{1}{\hat{H}-\varepsilon_v+i\omega}\hat{P}_c \right| \psi_v\Phi_\mu \right\rangle\right] \tag{3.34}$$

The computational cost is reduced even further by considering the following orthonormal basis $|t_\alpha\rangle$, which is used to expand the state $\hat{P}_c\,|\psi_v\Phi_\mu\rangle$ that spans the $N_v \cdot N_p$ linear space, we get

$$\hat{P}_c|\psi_v\Phi_\mu\rangle \approx \sum_\alpha^{N_T} |t_\alpha\rangle\langle t_\alpha|\hat{P}_c|\psi_v\Phi_\mu\rangle = \sum_\alpha^{N_T} |t_\alpha\rangle T_{\alpha\mu} \tag{3.35}$$

$$P^o_{\mu\nu}(i\omega) = -4\,\mathrm{Re}\left[\sum_{v,\alpha\beta} T^*_{\alpha,\nu\mu} \left\langle t_\alpha \left| \frac{1}{\hat{H}-\varepsilon_v+i\omega} \right| t_\beta \right\rangle T_{\beta,v\nu}\right] \tag{3.36}$$

Expanding the reducible dynamical polarizability in terms of the optimal basis

$$\Pi(\boldsymbol{r},\boldsymbol{r}';i\omega) \approx \sum_{\mu\nu} \Phi_\mu(\boldsymbol{r})\Pi_{\mu\nu}(i\omega)\Phi_\nu(\boldsymbol{r}') \tag{3.37}$$

where the matrix elements $\Pi_{\mu\nu}(i\omega)$ are given by:

$$\Pi_{\mu\nu}(i\omega) = \left[P(i\omega)\left[1-vP(i\omega)\right]^{-1}\right]_{\mu\nu} \tag{3.38}$$

where $v_{\mu\nu}$ is the frequency independent bare Coulomb interaction expanded in the same optimal polarizability basis $|\Phi_\mu\rangle$.

$$v_{\mu\nu} = \int d\mathbf{r}d\mathbf{r}'\langle\Phi_\mu|v(\mathbf{r},\mathbf{r}')|\Phi_\nu\rangle \tag{3.39}$$

In performing ab initio calculations within the quasiparticle approximation, the single-shot $G_0W_0$ method is frequently used. This method utilizes Kohn-Sham orbitals from Density Functional Theory (DFT) as a starting point for estimating quasiparticle energies. These energies are calculated by applying a first-order correction to the Kohn-Sham eigenvalues obtained from DFT. This correction is derived from the GW approximation for self-energy, which accounts for the many-body effects beyond those captured by DFT. Thus, quasiparticle energies are expressed as follows:

$$\varepsilon_n^{QP} = \varepsilon_n^{KS} + \langle\phi_n|\,\Sigma^{GW}(\varepsilon_n^{QP}) - V_{xc}^{DFT}\,|\phi_n\rangle \tag{3.40}$$

It is noteworthy to point out that the results of the one-shot $G_0W_0$ approximation are highly dependent on the choice of the exchange-correlation functional [65, 66]. This dependency is quite apparent from Eq.3.40. Consequently, the choice of the exchange-correlation functional sets a foundational benchmark that can significantly influence the perturbative adjustment of quasiparticle energies within the $G_0W_0$ framework.

## 3.5 Neutral Excitation

For charged excitation the one particle green's function is sufficient to obtain the ionization potential and electron affinity consequently the fundamental band gap is comparable with the results obtained from the photoemssion experiments. This however is not the case for neutral excitations, where an electron-hole pair is formed, such pair is referred to as **exciton**. What binds the electron-hole pair together is their mutual Coulomb interactions, describing such a two-body interaction using the one-particle green's function is not adequate.

The key quantity used to describe neutral excitation is the two-particle correlation function L. Many of it's properties are directly connected to experiments, most importantly the total energy can be expressed exactly in terms of L since the Coulomb interaction term is a two-body interaction [47].

One can define the two-particle correlation function as the functional derivative of the one-particle Green's function $G(1,1')$ with respect to a non-local external potential $U(2,2')$[67].

$$L(1,2;1',2') = \frac{\delta G(1,1')}{\delta U(2',2)} \tag{3.41}$$

The index (1) comprise position,time and spin coordinates: $(1) = (r_1, t_1, \sigma_1)$ similarly for the other indices. Making use of the identity App.A4 the correlation function can be re-written in terms of the inverse of the Green's function as follows:

$$L(1,2;1',2') = -\int d(34) G(1,3) \frac{\delta G^{-}1(3,4)}{\delta U(2',2)} G(4,1') \tag{3.42}$$

we can use the Dyson equation of the inverse one-particle Green's function which is given by:

$$G^{-1}(1,2) = G_0^{-1}(1,2) - \Sigma(1,2) - U(1,2) \tag{3.43}$$

Substituting Eq.3.43 into Eq.3.42 we get:

$$L(1,2;1',2') = -1 \int d(34)G(1,3)\frac{\delta\left[G_0^{-1}(3,4) - \Sigma(3,4) - U(3,4)\right]}{\delta U(2',2)}G(4,1') \quad (3.44)$$

$$= -1 \int d(34)G(1,3)\left[\frac{\delta G_0^{-1}}{\delta U(2',2)} - \delta(3,2')\delta(4,2) - \frac{\delta\Sigma(3,4)}{\delta U(2',2)}\right]G(4,1') \quad (3.45)$$

$$= G(1,2')G(2,1') - \int d(34)G(1,3)G(4,1')\left[\frac{\delta V_H(3)(34)}{\delta U(2',2)} + \frac{\delta\Sigma(3,4)}{\delta U(2',2)}\right] \quad (3.46)$$

using the chain rule $\frac{\delta\Sigma}{\delta U} = \left(\frac{\delta\Sigma}{\delta G}\right)\left(\frac{\delta G}{\delta U}\right)$ we get

$$L(1,2;1',2') = G(1,2')G(2,1') - \int d(3456)G(1,3)G(4,1')\left[\frac{\delta V_H(3)\delta(34)}{\delta G(5,6)} + \frac{\delta\Sigma(3,4)}{\delta G(5,6)}\right]\frac{\delta G(5,6)}{\delta U(2',2)} \quad (3.47)$$

The Hartree potential is given in terms of the bare Coulomb interaction v and the density $G(4,4^+)$ [68]:

$$V_H(3) = -i\int d4v(34)G(4,4^+) \quad (3.48)$$

We recover the expression given by Strinati [67]:

$$L(1,2;1',2') = G(1,2')G(2,1') + \int d(3456)G(1,3)G(4,1')\left[-iv(36)\delta(34)\delta(56) + \frac{\delta\Sigma(3,4)}{\delta G(6,5)}\right]\frac{\delta G(6,5)}{\delta U(2',2)} \quad (3.49)$$

re-writing it in a compact form we get:

$$L(1,2;1',2') = L_0(1,2;1',2') + \int d(3456)L_0(1,4;1,3')\Xi(3,5;4,6)L(6,2;5,2') \quad (3.50)$$

$L_0(1,2;1'2') = G(1,2')G(2,1')$ corresponds to a free electeon-hole pair with the interaction switched off. We also introduced the 4-point interaction kernel

$$\Xi(3,5;4,6) = -iv(36)\delta(34)\delta(56) + \frac{\delta\Sigma(3,4)}{\delta G(6,5)} \quad (3.51)$$

The kernel connects the independent-particle correlation function $L_0$ to the interacting L. This Dyson-like equation Eq.3.50 is known as the Bethe-Salpeter Equation (BSE) which was first developed in the framework of nuclear physics [69], then later on utilized in condensed matter physics to study neutral excitation and optical properties of semiconductors [70, 71]. In principle BSE is an exact equation Eq.3.50, solving it should give the two-particle correlation function which would give access to the properties of the neutral excitation. However, there are two major limitations; the exact form of the self energy $\Sigma$ is not known, even if it was, BSE as it stands is a 4-point problem. This means that one has to solve an equation with four space,time and spin arguments which is a major drawback for tractable calculations.

## 3.6 Bethe-Salpeter Equation In Practice

For optical excitation, one is mainly interested in the simultaneous creation and annihilation of the electron-hole pair. This means that instead of having four distinct time variables in Eq.3.50, two creations and two annihilations, we only focus on two. One instant where both electron-hole pair are created simultaneously, and another instant where both are annihilated simultaneously $L(1,2,(t_2=t_1);1',2',(t_{2'}=t_{1'}))$ . Consequently, and in the absence of external fields, only the time differences between the two-simultaneous processes is of interest [13]. A Fourier transform of the time difference to the frequency domain is carried out accordingly [72].

$$L(1,2;1',2') \Rightarrow L(1,2;1',2';\omega)$$

where the variable (1) indicates only the space and spin degrees of freedom.

Using the quasipaticle approximation for the one-body Green's function, the correlation

function with the interaction kernel switched off can be written as follows [13]:

$$L_0(1,2;1'2';\omega) = i\sum_{v,c}\left[\frac{\psi_c(\boldsymbol{x}_1)\psi_v^*(\boldsymbol{x}'_1)\psi_v(x_2)\psi_c^*(\boldsymbol{x}'_2)}{\omega-(\varepsilon_c-\varepsilon_v)+i\eta} - \frac{\psi_v(\boldsymbol{x}_1)\psi_c^*(\boldsymbol{x}'_1)\psi_c(\boldsymbol{x}_2)\psi_v^*(\boldsymbol{x}'_2)}{\omega+(\varepsilon_c-\varepsilon_v)-i\eta}\right] \tag{3.52}$$

where $\eta$ is a positive infinitesimal, v and c run over the occupied hole states and empty electron states respectively. In a similar manner, the correlation function $L(1,2;1',2';\omega)$ in the frequency domain can be written as follows:

$$L(1,2;1',2';\omega) = i\sum_{S}\left[\frac{\chi_S(\boldsymbol{x}_1,\boldsymbol{x}'_1)\chi_S{}^*(\boldsymbol{x}'_2,\boldsymbol{x}_2)}{\omega-\Omega_S} - \frac{\chi_S(\boldsymbol{x}_2,x'_2)\chi_S{}^*(\boldsymbol{x}'_1,\boldsymbol{x}_1)}{\omega+\Omega_S}\right] \tag{3.53}$$

where $\chi_S$ represent the electron-hole amplitudes, which have the following structure:

$$\chi_S(\boldsymbol{x},\boldsymbol{x}') = -\langle\Psi_0|\,\psi^\dagger(\boldsymbol{x}')\psi(\boldsymbol{x})\,|\Psi_S\rangle \tag{3.54}$$

Going to transition space, the electron-hole amplitudes can be written as:

$$\chi_S(\boldsymbol{x}',\boldsymbol{x}) = \sum_{v}^{occ}\sum_{c}^{empty} A_{vc}^S\psi_c(\boldsymbol{x})\psi_v^*(\boldsymbol{x}') + B_{vc}^S\psi_v(\boldsymbol{x})\psi_c^*(\boldsymbol{x}') \tag{3.55}$$

BSE can be written as an effective eigenvalue problem [13, 67]

$$(\varepsilon_c-\varepsilon_v)A_{vc}^S + \sum_{v'c'}K_{vc,v'c'}^{AA}(\Omega_S)A_{v'c'}^S + \sum_{v'c'}K_{vc,v'c'}^{AB}(\Omega_S)B_{v'c'}^S = \Omega_S A_{vc}^S \tag{3.56}$$

$$\sum_{v'c'}K_{vc,v'c'}^{BA}(\Omega_S)A_{v'c'}^S + (\varepsilon_c-\varepsilon_v)B_{vc}^S + \sum_{v'c'}K_{vc,v'c'}^{BB}(\Omega_S)B_{v'c'}^S = -\Omega_S B_{vc}^S \tag{3.57}$$

The BSE Hamiltonian has the structure

$$H^{BSE}(\Omega_S) = \begin{pmatrix}(\varepsilon_c-\varepsilon_v) & 0\\ 0 & (\varepsilon_v-\varepsilon_c)\end{pmatrix} + \begin{pmatrix}K^{AA}(\Omega_S) & K^{AB}(\Omega_S)\\ K^{BA}(\Omega_S) & K^{BB}(\Omega_S)\end{pmatrix} \tag{3.58}$$

where elements of the electron-hole interaction kernel are calculated in the basis of

the single-particle orbitals. The diagonal elements have the following form

$$K^{AA}_{vc,v'c'}(\Omega_S) = i \int d(3456)\psi_v(\boldsymbol{x}_4)\psi^*_c(\boldsymbol{x}_3)\Xi(35,46;\Omega_S)\psi^*_{v'}(\boldsymbol{x}_5)\psi_{c'}(\boldsymbol{x}_6) \tag{3.59}$$

$$K^{BB}_{vc,v'c'}(\Omega_S) = i \int d(3456)\psi_v(\boldsymbol{x}_3)\psi^*_c(\boldsymbol{x}_4)\Xi(35,46;\Omega_S)\psi^*_{v'}(\boldsymbol{x}_6)\psi_{c'}(\boldsymbol{x}_5) \tag{3.60}$$

The off-diagonal matrix elements

$$K^{AB}_{vc,v'c'}(\Omega_S) = i \int d(3456)\psi_v(\boldsymbol{x}_4)\psi^*_c(\boldsymbol{x}_3)\Xi(35,46;\Omega_S)\psi^*_{v'}(\boldsymbol{x}_6)\psi_{c'}(\boldsymbol{x}_5) \tag{3.61}$$

$$K^{BA}_{vc,v'c'}(\Omega_S) = i \int d(3456)\psi_v(\boldsymbol{x}_3)\psi^*_c(\boldsymbol{x}_4)\Xi(35,46;\Omega_S)\psi^*_{v'}(\boldsymbol{x}_5)\psi_{c'}(\boldsymbol{x}_6) \tag{3.62}$$

### 3.6.1 Tamm-Dancoff Approximation

In principle, we can obtain the excitation spectra by solving the effective BSE eigenvalue problem Eq.3.63. However, this is not an easy task as the equation is non-Hermitian. One way to bypass this hurdle is to neglect the coupling between the excitation and the de-excitation process. This is done by neglecting the off-diagonal elements in the BSE Hamiltonian Eq.3.58, which has negligible effects on the excitation energies [73, 13]. As a result, Hermiticity is enforced. This is valid as long as the energy of the electron-hole interaction is small compared to the quasi-particle gap. Setting the off-diagonal terms in the BSE Hamiltonian Eq.3.58 to zero, separates it into two block-diagonal parts, the resonant and anti-resonant contributions, which are active for positive and negative frequencies respectively. We consider the resonant transitions, consequently we need to solve the following eigenvalue problem

$$\sum_{v'c'} H^{BSE}_{vc,v'c'} A^S_{v'c'} = \Omega_S A^S_{vc} \tag{3.63}$$

where, v and c run over the occupied hole (valance) states and empty electron (conduction) states respectively, $A^S_{vc}$ are the electron-hole amplitudes, $\Omega_S$ is the excitation energy and $H^{BSE}_{vc,v'c'}$ is the effective electron-hole Hamiltonian given by:

$$H^{BSE}_{vc,v'c'} = (\varepsilon_c - \varepsilon_v)\delta_{vv'}\delta_{cc'} + K^{AA}_{vc,v'c'} \tag{3.64}$$

where $K^{AA}_{vc,v'c'}$ are the matrix elements for the BSE kernel for resonant transitions given by Eq.3.59. Solving the effective eigenvalue problem Eq.3.63 one then obtains the excitation energies $\Omega_s$ which is then added to the ground state Energy $E_0$ to get the excited state energy $E_S$

$$E_S = E_0 + \Omega_S \tag{3.65}$$

Neglecting the off-diagonal terms in the Hamiltonian and considering only the resonant transitions leads to the Tamm-Dancoff approximation, hence why Eq.3.63 is often refereed to as the Tamm-Dancoff approximation to BSE, **TDA-BSE** [15].

The many-body excited state $|\Psi_S\rangle$ can be constructed from the ground state $|\Psi_0\rangle$ as follows

$$|\Psi_S\rangle = \sum_v \sum_c A^S_{vc} \hat{a}_v \hat{a}^\dagger_c |\Psi_0\rangle \tag{3.66}$$

$$|\Psi_S\rangle = \sum_v \sum_c A^S_{vc} |vc\rangle \tag{3.67}$$

where applying $\hat{a}_v$,$\hat{a}^\dagger_c$ to the many electron ground state annihilates or creates an electron, respectively. The many electron ground state is constructed from a single

Slater determinant of $N_v$ single particle Kohn-Sham states:

$$|\Psi_0\rangle = \frac{1}{\sqrt{N_v}} det|\psi_1 ... \psi_{N_v}| \tag{3.68}$$

which in the second quantization framework can be written as:

$$|\Psi_0\rangle = \frac{1}{\sqrt{N_v!}} \left( \prod_{\nu} \hat{a}_{\nu}^{\dagger} \right) |0\rangle \tag{3.69}$$

### 3.6.2 GW@BSE

The standard approximation for the self energy within BSE is the GW approximation [13, 71, 67, 74]. This approximation, proved to be successful in describing gas-phase molecules [75], solids [76]. In the GW approximation, the self energy is given by $\Sigma = iGW$ to the first order in W, plugging it into the BSE kernel. Taking the derivative with respect to G, we neglect the higher order perturbation in W [72, 13]. The BSE kernel reads:

$$\Xi(3,5;4,6) = -iv(36)\delta(34)\delta(56) + iW(3^{+},4)\delta(36)\delta(45) \tag{3.70}$$

The first term is the bare Coulomb interaction $v$ which represent the repulsive exchange interactions ($K^x$). It governs the properties of the excitation spectrum, such as the splitting between spin-singlet and spin-triplet excitation [13]. The second term, on the other hand, is the screened Coulomb interaction W which represents the electron-hole attraction, also known as the direct term ($K^d$) [13, 72]. The BSE kernel given by Eq.3.70 can be re-written in an abbreviated way as:

$$\Xi(3,5;4,6) = K^x(35;46) + K^d(35;46) \tag{3.71}$$

The current form of the BSE kernel is fully dynamical, we can make further approximation to simplify the calculations by taking the screening interaction W to be static [70, 71].

$$W(r_1, r_2, \omega) \Rightarrow W(r_1, r_2, \omega = 0) \tag{3.72}$$

The use of the static approximation is widely spread [77], mainly because the quasiparticle re-normalization effects cancel out the dynamical effects in the BSE kernel [78]. Therefore, neglecting dynamical effects should yield accurate results. The BSE Hamiltonian can be written as:

$$H^{BSE}_{vc,v'c'} = \hat{D}_{vv'cc'} + \hat{K}^{AA,x}_{vc,v'c'} + \hat{K}^{AA,d}_{vc,v'c'} \tag{3.73}$$

where $\hat{D}$ is the $(\varepsilon_c - \varepsilon_v)\delta_{vv'}\delta_{cc'}$ term in Eq.3.64 often referred to as the diagonal term. The matrix elements for the exchange interaction $K^x$ and direct interactions $K^d$ using the electron-hole states $|vc\rangle$ can be written as:

$$\langle vc|K^{AA,x}|v'c'\rangle = \int d\mathbf{r}d\mathbf{r}'\psi^*_c(\mathbf{r})\psi_v(\mathbf{r})v(\mathbf{r},\mathbf{r}')\psi_{v'}(\mathbf{r}')\psi^*_{c'}(\mathbf{r}') \tag{3.74}$$

$$\langle vc|K^{AA,d}|v'c'\rangle = -\int d\mathbf{r}\, d\mathbf{r}'\, \psi^*_c(\mathbf{r})\psi_{c'}(\mathbf{r})\psi_v(\mathbf{r}')\psi^*_{v'}(\mathbf{r}')W(\mathbf{r},\mathbf{r}';\omega = 0) \tag{3.75}$$

the direct interaction term in the static limit, which fully ignores the dynamical properties of W.

One thing which hasn't been mentioned thus far, is the spin degrees of freedom. The spin structure of the solutions of the BSE Eq.3.63 is highly dependent on the strength of the spin-orbit interactions [13]. In the case of weak spin-orbit coupling, which can be neglected. The Hamiltonian can be written in a block diagonal form using the singlet and triplet pair states. As a result the singlet and triplet states can be treated independently [79]. The effective eigenvalue problem needs to be solved

for two separate Hamiltonian pairs

$$\hat{H}^{BSE}_{singlet} = \hat{D} + 2\hat{K}^x + \hat{K}^d \tag{3.76}$$

$$\hat{H}^{BSE}_{triplet} = \hat{D} + \hat{K}^d \tag{3.77}$$

$K^x$ and $K^d$ are the exchange and direct interactions respectively, their matrix elements are given by Eq.3.74, Eq.3.75. The work presented in this thesis deals with the case of weak spin-orbit interactions, which is the common case [80].

### 3.6.3 Breaking The Curse Of Dimensionality

The eigenvalue problem Eq.3.63 posed by the BSE is numerically cumbersome, which is a huge drawback of BSE calculations. A single excitoninc state is described by a vector of dimension $N_v \times N_c$, which means that the excitonic state scales as $N^2$ with the size of the system indicated as N. The BSE Hamiltonian on the other hand can described by a $(N_v \times N_c)$ cubic matrix. Applying the BSE Hamiltonian on an excitonic state scales as $N^4$, such scaling hinders the use of BSE to study large model structures. Marsili et.al [80] introduced a new approach to BSE calculations which reduces the scaling to $N^3$ with respect to system size, making it more computationally efficient. They utilized the localization properties of Wannier's functions [81]. In this part we give a quick overview of the approach they followed .

There are three important ingredients used to break the curse of $N^4$ scaling. First, define a new set of functions known as batch representation [82, 83] instead of using the electron-hole product state $|vc\rangle$.

$$\xi_\nu(x) = \sum_c A_{vc}\psi_c(x) \tag{3.78}$$

where $A_{vc} = \langle \psi_c | \xi_v \rangle$ and $|\psi\rangle$ are the eigenstates of the Kohn-Sham Hamiltonian

$$\hat{H}_{KS} |\psi_i\rangle = \varepsilon_i |\psi_i\rangle \tag{3.79}$$

We note that these new functions are restricted to the manifold of empty KS states. In this representation a generic excitonic state in the Tamm-Dancoff given by Eq.3.66 can be represented in terms of $N_v \xi_v$ as:

$$|\Psi_S\rangle = \sum_v \left[ \hat{a}_v \sum_c \langle \psi_c | \xi_v \rangle \hat{a}_c^\dagger \right] |\Psi_0\rangle \tag{3.80}$$

We now summarize the action of the different terms of $\hat{H}^{BSE}$ for different spin configurations Eq.3.76-3.77 on $\{\xi_v(\mathbf{r})\}$ and $\{\tilde{\xi}_v(\mathbf{r})\}$ where the curly brackets in $\{\xi_v\}$ indicates the entire set of functions for all possible values of v.

A general case of system with time-reversal symmetry is considered, therefore the one-body states can be taken to be real, hence why the complex-conjugate sign is dropped from the following equations. The action of the diagonal term on the batch representation is given by:

$$|\{\xi'\}\rangle = \hat{D} |\{\xi\}\rangle \tag{3.81}$$

$$|\xi'_\nu\rangle = (\hat{H}^{qp} - \varepsilon_v I) |\xi_\nu\rangle \tag{3.82}$$

where $H^{qp}$ is the quasi particle Hamiltonian, which when using the diagonal form of GW its egienfunctions coincide with the ones of the Kohn-Sham states. The computational cost of applying the $\hat{D}$ operator on a single $\xi_v$ scales as $N \log N$, for $N_v$ valance states, it will scale as $N^2 \log N$.

The second term in the BSE Hamiltonian is the exchange operator, its action on $\{\xi_v\}$ can be written as:

$$|\{\xi''\}\rangle = \hat{K}^x |\{\xi\}\rangle \tag{3.83}$$

$$\xi''_v(\mathbf{r}) = \int P_c(\mathbf{r},\mathbf{r}')\psi_v(\mathbf{r}')v(\mathbf{r}',\mathbf{r}'')\psi_{v'}(\mathbf{r}'')\xi_{v'}(\mathbf{r}'')d\mathbf{r}'d\mathbf{r}'' \tag{3.84}$$

where $\hat{P}_c$ the projector onto the KS unoccupied manifold, the second ingredient in breaking the $N^4$ scaling,

$$P_c(\mathbf{r},\mathbf{r}') = \sum_c \psi_c(\mathbf{r})\psi_c(\mathbf{r}') = \delta(\mathbf{r}-\mathbf{r}') - \sum_v \psi_v(\mathbf{r})\psi_v(\mathbf{r}') \tag{3.85}$$

Note that by using the projection operator Eq.3.84 doesn't require any calculations over unoccupied states. As a result the computational cost of applying $\hat{K}^x$ to $|\{\xi\}\rangle$ scales as $N^3$.

The Final term in the BSE Hamiltonian is the direct term, which we can write its effect on the batch representation as:

$$|\{\xi'''\}\rangle = \hat{K}^d|\{\xi\}\rangle \tag{3.86}$$

$$\xi'''_\nu(\mathbf{r}) = -\int P_c(\mathbf{r},\mathbf{r}')\sum_{v'}\xi_{v'}(\mathbf{r}')\left[\int W(\mathbf{r}',\mathbf{r}'')\psi_{\nu'}(\mathbf{r}'')\psi_\nu(\mathbf{r}'')d\mathbf{r}''\right]d\mathbf{r}' \tag{3.87}$$

The computational cost of applying $\hat{K}^d$ on a generic state $|\{\xi\}\rangle$ scales as $N^4$, working with the batch representation is not enough to reduce the scaling, here is where the third ingredient comes into play. Instead of working with the Kohn-Sham representation, the valance states are transformed to maximally localized Wannier's functions $w_v$ using the unitary transformation matrix U[81].

$$w_v(\mathbf{r}) = \sum_{v'} U_{vv'}\psi_{v'}(\mathbf{r}) \tag{3.88}$$

Now, the function $\xi_v$ given in Eq.3.78 can be expressed as:

$$\tilde{\xi}_v(\mathbf{x}) = \sum_{v'} U_{vv'}\xi_{v'}(\mathbf{x}) \tag{3.89}$$

Using the Wannier's representation:

$$|\{\tilde{\xi}'''\}\rangle = \hat{K}^d|\{\tilde{\xi}\}\rangle \tag{3.90}$$

$$\tilde{\xi}'''_\nu(\mathbf{r}) = -\int P_c(\mathbf{r},\mathbf{r}')\sum_{v'}\tilde{\xi}_{v'}(\mathbf{r}')\left[\int W(\mathbf{r}',\mathbf{r}'')w_{\nu'}(\mathbf{r}'')w_\nu(\mathbf{r}'')d\mathbf{r}''\right]d\mathbf{r}' \tag{3.91}$$

By considering only overlapping pairs of Wannier's functions and ignoring terms involving products of non overlapping ones according to a user defined threshold s

$$\int |w_v(\mathbf{r})|^2|w_{v'}(\mathbf{r})|^2 d\mathbf{r} > s \tag{3.92}$$

The computational cost is reduced to $N^3$.

In practice, when evaluating the $\hat{K}^d$ it is decomposed into a bare $\hat{K}^{d,b}$ and a correlation part $\hat{K}^{d,c}$ by separating the screened Coulomb interaction into: the bare Coulomb potential $v(\mathbf{r},\mathbf{r}')$ and the correlation part of the screened Coulomb potential $W_c(\mathbf{r},\mathbf{r}')$

$$W(\mathbf{r},\mathbf{r}') = v(\mathbf{r},\mathbf{r}') + W_c(\mathbf{r},\mathbf{r}') \tag{3.93}$$

The bare part of direct interaction $\hat{K}^d$ acting on $|\{\tilde{\xi}\}\rangle$

$$\left|\{\tilde{\xi}'''^{,b}\}\right\rangle = \hat{K}^{d,b}\left|\{\tilde{\xi}\}\right\rangle \tag{3.94}$$

$$\tilde{\xi}'''^{,b}_v(\mathbf{r}) = -\int P_c(\mathbf{r},\mathbf{r}')\sum_{v'}\tau^b_{vv'}(\mathbf{r}')\tilde{\xi}_{v'}(\mathbf{r}')d\mathbf{r}' \tag{3.95}$$

with $\tau^b_{vv'}(\mathbf{r})$ being evaluated for the overlapping Wannier's functions according to user defined threshold Eq.3.92

$$\tau^b_{vv'}(\mathbf{r}) = \int v(\mathbf{r},\mathbf{r}')w_{v'}(\mathbf{r}')w_v(\mathbf{r}')d\mathbf{r}' \tag{3.96}$$

The action of the correlation part of $\hat{K}^{d,c}$

$$|\{\tilde{\xi}^{\prime\prime\prime,c}\}\rangle = \hat{K}^{d,c}|\{\tilde{\xi}\}\rangle \tag{3.97}$$

$$\tilde{\xi}^{\prime\prime\prime,c}_{v}(\mathbf{r}) = -\int P_c(\mathbf{r},\mathbf{r}')\tau^{c}_{vv'}(\mathbf{r}')\tilde{\xi}_{v'}(\mathbf{r}')d\mathbf{r}' \tag{3.98}$$

with

$$\tilde{\tau}^{c}_{vv'}(\mathbf{r}) = \int W_c(\mathbf{r},\mathbf{r}')w_{v'}(\mathbf{r}')w_{v}(\mathbf{r}')d\mathbf{r}' \tag{3.99}$$

The screened Coulomb interaction can be expressed using the optimal basis set as follows,

$$W_c(\mathbf{r},\mathbf{r}') = \sum_{\mu\nu}(v\Phi_\mu)(\mathbf{r})\Pi_{\mu\nu}(v\Phi_\nu)(\mathbf{r}') \tag{3.100}$$

The matrix elements $\Pi_{\mu\nu}(i\omega)$ are given by Eq.3.37. Using the notation

$$(v\Phi_\mu)(\mathbf{r}) = \int d\mathbf{r}' v(\mathbf{r},\mathbf{r}')\Phi_\mu(\mathbf{r}') \tag{3.101}$$

$\tau^{c}_{vv'}$ Eq.3.99 can be expressed as

$$\tau^{c}_{vv'}(\mathbf{r}) = \sum_{\mu}(v\Phi_\mu)(\mathbf{r})Z^{\nu}_{v\mu} \tag{3.102}$$

The coefficients $Z^{\nu}_{v'\mu}$ are defined as:

$$Z^{\nu}_{v'\mu} = \sum_{\nu}\Pi_{\mu\nu}\int d\mathbf{r}(v\Phi_\nu)(\mathbf{r})w_\nu(\mathbf{r})w_{\nu'}(\mathbf{r}) \tag{3.103}$$

Now we have all the important ingredients to solve the Tamm-Dancoff-BSE and obtain the electron-hole excitation.

# Chapter Four

# Excited State Forces and Structural Optimization

## 4.1 Introduction

While the Bethe-Salpeter equation has advanced our understanding of neutral excitation and their effects on optical spectra, when it comes to capturing the dynamics of atomic positions and forces in excited states, it leaves a lot to be desired [84]. To this date only one pioneering study of the first derivative of the energy with respect to the nuclear displacements within the GW-BSE formalism have been published [84, 85].

In ground state electronic structure calculations, the forces acting on the $I^{th}$ nucleus are calculated by utilizing the Hellman-Feynman theorem [86, 87]. They can be written as [88]:

$$F_I = -\int n_R(\mathbf{r}) \frac{\partial V_R(\mathbf{r})}{\partial \mathbf{R}_I} d\mathbf{r} - \frac{\partial E_N(\mathbf{R})}{\partial \mathbf{R}_I} \tag{4.1}$$

where $n(\mathbf{r})$ is the ground-state electron charge density corresponding to the nuclear configuration ($\mathbf{R}$), $V_R(\mathbf{r})$ is the electron-nucleus interaction and $E_N(\mathbf{R})$ is the electrostatic interaction between different nuclei. These forces are then used to preform structural optimization, with the equilibrium condition being:

$$F_I \equiv -\frac{\partial E(\mathbf{R})}{\partial \mathbf{R}_I} = 0 \tag{4.2}$$

Forces obtained through Eq.4.1 are considered to be cheap byproduct of ground state electronic structure calculations. This, however is not the case for excited state calculations. The goal is to compute the forces acting on $I^{th}$ nucleus in the excited state using the Hellman-Feynman theorem:

$$\vec{F}_I = -\frac{\partial E_S}{\partial \mathbf{R}_I} = -\langle \Psi | \frac{\partial \hat{H}}{\partial \mathbf{R}_I} | \Psi \rangle \tag{4.3}$$

The energy of the excited states $E_S$ is given as the sum of the ground state energy $E_0(R)$ obtained from DFT calculations and the excitation energy $\Omega_S$ obtained from

solving the Bethe-Salpeter equation Eq.3.64.

$$E_S(\boldsymbol{R}) = E_0(\boldsymbol{R}) + \Omega_S(\boldsymbol{R}) \tag{4.4}$$

The forces are given by:

$$-\vec{F}_I = \frac{\partial E_S(\boldsymbol{R})}{\partial \boldsymbol{R}} = \frac{\partial E_0(\boldsymbol{R})}{\partial \boldsymbol{R}} + \frac{\partial \Omega_S(\boldsymbol{R})}{\partial \boldsymbol{R}} \tag{4.5}$$

The excitonic part reads:

$$\begin{aligned} -F &= \partial_R \Omega_S && (4.6)\\ &= \partial_R \langle \Phi | \, H_{BSE} \, | \Phi \rangle && (4.7)\\ &= \left[ \langle \partial_R \Phi | \Phi \rangle + \langle \Phi | \partial_R \Phi \rangle \right] \Omega_S + \langle \Phi | \, \partial_R H^{exc}_{BSE} \, | \Phi \rangle && (4.8)\\ &= \partial_R \langle \Phi | \Phi \rangle \, \Omega_S + \langle \Phi | \, \partial_R H_{BSE} \, | \Phi \rangle && (4.9)\\ &= \langle \Phi | \, \partial_R H_{BSE} \, | \Phi \rangle && (4.10) \end{aligned}$$

Note that the force is equal to the the derivative of the Hamiltonian, the variations of the wave function $\partial_R \Phi$ vanishes since we are using the exact eigen-functions of the BSE Hamiltonian given by Eq.[4.35]. Now, we can write Hellmann-Forces acting on the Ith nucleus in excited states:

$$F_I = - \langle \Phi | \, \partial_R H_{BSE} \, | \Phi \rangle + F_0 \tag{4.11}$$

where $F_0$ is the force in the ground state. As we mentioned in the previous chapter the BSE Hamiltonian is composed of three terms $\hat{D}, \hat{K}^x$ and $\hat{K}^d$ which we need to

take the derivative of with respect to the nuclear positions R.

$$\partial_R \Omega_S = \langle \Phi | \, \partial_R H^{exc}_{BSE} \, | \Phi \rangle \tag{4.12}$$

$$= \langle \Phi | \, \partial_R \hat{D}_{vc,v'c'} + \partial_R \hat{K}^x_{vc,v'c'} + \partial_R \hat{K}^d_{vc,v'c'} \, | \Phi \rangle \tag{4.13}$$

We will take the derivative of each term separately. The diagonal operator $\hat{D}$ is defined as :

$$\hat{D}_{vv'cc'} = (\epsilon_c - \epsilon_v)\delta_{vv'}\delta_{cc'}$$

The action of the diagonal term on the batch representation is given by:

$$|\xi'_v\rangle = (H^{KS} - \epsilon_v) \, |\xi_v\rangle \tag{4.14}$$

Hence, we have

$$\langle \Phi | \, \partial_R \hat{D} \, | \Phi \rangle = \langle \Phi | \, \partial_R (H^{KS} - \epsilon_v) \, | \Phi \rangle \tag{4.15}$$

$$= \langle \Phi | \, \partial_R H^{KS} - \partial_R \epsilon_v \, | \Phi \rangle \tag{4.16}$$

$$= \langle \Phi | \, \frac{H^{KS}_{\vec{R}+\delta} - H^{KS}_{\vec{R}-\delta}}{2\delta} - \partial_R \epsilon_v) \, | \Phi \rangle \tag{4.17}$$

where we used the finite difference method to obtain the derivative of the Kohn-Sham Hamiltonian.

The exchange term $\hat{K}^x$ is defined as

$$K^x_{vv'cc'} = \int\int dx dx' \psi^*_v(x)\psi_c(x) v_c(r,r')\psi_{v'}(x')\psi^*_{c'}(x') \tag{4.18}$$

The final term which is the direct screened coulomb potential:

$$K^d_{vv'cc'} = -\int\int dx dx' \psi^*_v(x)\psi_{v'}(x) W(r,r')\psi_c(x')\psi_c(x')\psi^*_{c'}(x) \tag{4.19}$$

where $\psi_i$ is the single particle state and $x = (r, \sigma)$ is a combined space and spin coordinate and $v_c$, W are the bare and screened Coulomb operators respectively. Note that the exchange term involves the Coulomb interactions, while the direct term involves only the screened Coulomb interaction. The direct term is the one responsible for the attractive nature of the electron-hole interaction, which leads to the formation of the bound electron-hole states , i.e excitons [13]. Since we are only dealing with real wave-functions we can drop the complex conjugate sign.

Utilizing the batch representation $|\{\xi\}\rangle$, the operator $\hat{K}^x$ acting on $|\{\xi\}\rangle$ can be written as

$$\left|\xi_v^{''}\right\rangle = K^x_{vv'} \left|\xi_v\right\rangle$$

where;

$$\left\langle x \middle| \xi_v^{''} \right\rangle = \int P_c(x, x')\psi_v(x')v_c(x', x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \tag{4.20}$$

we have

$$\langle\Phi| K^x_{vv'} |\Phi\rangle = \sum_{vv'} \int \xi_v(x)P_c(x, x')\psi_v(x')v_c(x', x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \tag{4.21}$$

where $P_c$ is the projector operator over the manifold of unoccupied states

$$P_c(x, x') = \sum_c \psi_c(x)\psi_c(x') \tag{4.22}$$

$$= \delta(x - x') - \sum_v \psi_v(x)\psi_v(x') \tag{4.23}$$

Now we want to evaluate the derivative of Eq.[4.21] with respect to R.

$$\begin{aligned}\langle\Phi|\,\partial_R K^x_{vv'}\,|\Phi\rangle =& \sum_{vv'}\int \xi_v(x)\partial_R P_c(x,x')\psi_v(x')v_c(x',x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \\ &+\sum_{vv'}\int \xi_v(x)P_c(x,x')\partial_R\psi_v(x')v_c(x',x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \\ &+\sum_{vv'}\int \xi_v(x)P_c(x,x')\psi_v(x')v_c(x',x'')\partial_R\psi_{v'}(x'')\xi_{v'}(x'')dx'dx''\end{aligned} \tag{4.24}$$

Now the first term we have the derivative of the projection operator which is given by:

$$\partial_R P_c = -\sum_{v'}|\partial_R\psi_{v'}\rangle\langle\psi_{v'}| - \sum_{v'}|\psi_{v'}\rangle\langle\partial_R\psi_{v'}| \tag{4.25}$$

The derivative of the projection operator acting on $\psi_v$ gives:

$$\partial_R P_c\,|\psi_v\rangle = -\sum_{v'}|\partial_R\psi_{v'}\rangle\,\delta_{vv'} - \sum_{v'}|\psi_{v'}\rangle\langle\partial_R\psi_{v'}|\psi_v\rangle \tag{4.26}$$

So we have

$$\begin{aligned}\langle\Phi|\,\partial_R K^x_{vv'}\,|\Phi\rangle =& -\sum_{vv'}\int \xi_v(x)\partial_R\psi_v(x)\psi_v(x')v_c(x',x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \\ &-\sum_{vv'}\int \xi_v(x)\psi_v(x)\partial_R\psi_{v'}(x')\psi_v(x')v_c(x',x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \\ &+\sum_{vv'}\int \xi_v(x)P_c(x,x')\partial_R\psi_v(x')v_c(x',x'')\psi_{v'}(x'')\xi_{v'}(x'')dx'dx'' \\ &+\sum_{vv'}\int \xi_v(x)P_c(x,x')\psi_v(x')v_c(x',x'')\partial_R\psi_{v'}(x'')\xi_{v'}(x'')dx'dx''\end{aligned} \tag{4.27}$$

Now the Final term is the direct term Eq.[4.19], which we can write its effect on the batch representation as:

$$\left|\{\xi'''_v\}\right\rangle = K^d_{vv'}\,|\{\xi_v\}\rangle$$

where;

$$\left\langle x\middle|\xi_v^{'''}\right\rangle = -\int P_c(x,x')\sum_{v'}\xi_{v'}(x')\left[\int W(x',x'')\psi_{v'}(x'')\psi_v(x'')dx''\right]dx' \tag{4.28}$$

still the computational cost of evaluating the direct term is high, that is why we use the Wannier's representation for the excited states [80].

$$\left|\{\xi_v^{'''}\}\right\rangle = K^d_{vv'}\left|\{\tilde{\xi}_v\}\right\rangle$$

$$\left\langle x\middle|\xi_v^{'''}\right\rangle = -\int P_c(x,x')\sum_{v'}\tilde{\xi}_{v'}(x')\left[\int W(x',x'')w_{v'}(x'')w_v(x'')dx''\right]dx' \tag{4.29}$$

so we have:

$$\begin{aligned}\langle\Phi|\,\partial_R K^d_{vv'}\,|\Phi\rangle = &-\sum_{vv'}\int\tilde{\xi}_v(x')\partial_R P_c(x,x')\tilde{\xi}_{v'}(x')\left[\int W(x',x'')w_{v'}(x'')w_v(x'')dx''\right]dx'\\ &-\sum_{vv'}\int\tilde{\xi}_v(x')P_c(x,x')\tilde{\xi}_{v'}(x')\left[\int W(x',x'')\partial_R w_{v'}(x'')w_v(x'')dx''\right]dx'\\ &-\sum_{vv'}\int\tilde{\xi}_v(x')P_c(x,x')\tilde{\xi}_{v'}(x')\left[\int W(x',x'')w_{v'}(x'')\partial_R w_v(x'')dx''\right]dx'\\ &-\sum_{vv'}\int\tilde{\xi}v(x')P_c(x,x')\tilde{\xi}v'(x')\left[\int \partial_R W(x',x'')w_{v'}(x'')w_v(x'')dx''\right]dx'\end{aligned} \tag{4.30}$$

where the projection operator using the Wannier's representation is written as:

$$P_c(x,x') = \sum_c w_c(x)w_c(x') \tag{4.31}$$

$$= \delta(x-x') - \sum_v w_v(x)w_v(x') \tag{4.32}$$

The derivative of The projection operator acting on $w_v$ gives:

$$\partial_R P_c\,|w_v\rangle = -\sum_{v'}|\partial_R w_{v'}\rangle\,\delta_{vv'} - \sum_{v'}|w_{v'}\rangle\,\langle\partial_R w_{v'}|w_v\rangle \tag{4.33}$$

Implementing these equations to obtain the forces is not easy by any means. Hence why we explore another approach.

In the next section we explore a way to calculate the excited state forces within the Tamm-Dancoff approximation to BSE (TDA-BSE) that we discussed briefly in Sec.3.6.1.

## 4.2 Forces within TDA-BSE

A generic state of a many-body system, encompassing the full spectrum of behaviors under various conditions, can be expressed as:

$$|\Psi\rangle = \alpha\,|\Psi_0\rangle + |\Psi_S\rangle \tag{4.34}$$

Here, $|\Psi_0\rangle$ represents the many-body ground state, as defined by Eq.4.41 within the framework of second quantization formalism. This ground state is scaled by an overall factor $\alpha$ that governs it's contribution to the overall state $|\Psi\rangle$. $|\Psi_S\rangle$ denotes the collection of excited states, which are un-normalized in this context.

Under the Tamm-Dancoff Approximation (TDA) [3.6.1] the excited state $|\Psi_S\rangle$ is given by Eq.3.66. Adapting the batches representation as in Marsili et.al [80] [3.6.3], the excited states $|\Psi_S\rangle$ are constructed as follows:

$$|\Psi_S\rangle = \begin{pmatrix} \xi_1 \\ .. \\ .. \\ .. \\ \xi_{N_v} \end{pmatrix} \tag{4.35}$$

Here, each $\xi_v$ is defined as

$$|\xi_v\rangle = \sum_c A_{vc} |\psi_c\rangle \tag{4.36}$$

with $A_{v,c} = \langle\psi_c|\xi_v\rangle$ and $|\psi\rangle$ are the eigenstates of the Kohn-Sham Hamiltonian

$$\hat{H}_{KS} |\psi_i\rangle = \varepsilon_i |\psi_i\rangle \tag{4.37}$$

It is safe to say that the core components of our excited state wave functions are the one-body KS wavefunctions, denoted as $\psi_i$. To explore the impact of a perturbation on these wavefunctions, let's consider a scenario where a small perturbation is introduced to the KS Hamiltonian, $\hat{H}'_{KS}$:

$$\hat{H}'_{KS} |\psi'_i\rangle = \varepsilon'_i |\psi'_i\rangle \tag{4.38}$$

where $|\psi'_i\rangle$ denotes the perturbed states and $\varepsilon'_i$ is the corresponding energy of these perturbed states.

Within the Tamm-Dancoff approximation, a subspace is constructed using the perturbed one-body Kohn-Sham wave functions. This allows for a projection of the unperturbed excited state into the perturbed manifold as follows:

$$|\tilde{\Psi}_S\rangle = \hat{\mathcal{P}}'|\Psi_S\rangle = \sum_{v,c} \hat{a}'^{\dagger}_c \hat{a}'_v |\Psi'_0\rangle\langle\Psi'_0|\hat{a}'^{\dagger}_v \hat{a}'_c|\Psi_S\rangle \tag{4.39}$$

This projection captures how the unperturbed excited states overlap with the new states defined in the perturbed system. $|\Psi'_0\rangle$ is the ground state of the perturbed system, which can be constructed as a Slater determinant of the perturbed wavefunctions:

$$|\Psi'_0\rangle = \frac{1}{\sqrt{N_v}} det|\psi'_1 ... \psi'_{N_v}| \tag{4.40}$$

In the second quantization for:

$$|\Psi'_0\rangle = \frac{1}{\sqrt{N_v!}} \left( \prod_{\nu} \hat{a}'^{\dagger}_{\nu} \right) |0\rangle \tag{4.41}$$

Recall the unperturbed excited state is given by:

$$|\Psi_S\rangle = \sum_{v,c} A_{v,c} \hat{a}^{\dagger}_c \hat{a}_v \, |\Psi_0\rangle \tag{4.42}$$

The un-normalised coefficient $\tilde{A}_{v',c'}$ for the perturbed system become

$$\tilde{A}_{v',c'} = \sum_{v,c} A_{v,c} \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{c'} \psi'_{v'+1} \dots \psi'_{N_v} | \psi_1 \dots \psi_{v-1} \psi_c \psi_{v+1} \dots \psi_{N_v} \rangle \right| \tag{4.43}$$

Using the batches for the unperturbed Hamiltonian given in Eq.4.48 $\tilde{A}_{v',c'}$ reads:

$$\tilde{A}_{v',c'} = \sum_{v} \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{c'} \psi'_{v'+1} \dots \psi'_{N_v} | \psi_1 \dots \psi_{v-1} \xi_v \psi_{v+1} \dots \psi_{N_v} \rangle \right| \tag{4.44}$$

Manipulating the determinant to reposition $\psi'_{c'}$ to the first row. The antisymmetry property dictates that swapping any two rows changes the sign of the determinant. As a result we get:

$$\tilde{A}_{v',c'} = \sum_{v,c} A_{v,c} (-1)^{v-1} \det \left| \langle \psi'_{c'} \psi'_1 \dots \psi'_{v'-1} \psi'_{v'+1} \dots \psi'_{N_v} | \psi_1 \dots \psi_{v-1} \xi_v \psi_{v+1} \dots \psi_{N_v} \rangle \right| \tag{4.45}$$

Using the co-factor expansion along the row containing $\xi_v$ of the we get:

$$\begin{aligned} \tilde{A}_{v',c'} = \sum_{v} (-1)^{v'-1} \times \Bigg[ & \sum_{v'' \neq v} (-1)^{v''-1} \\ & \langle \psi'_{c'} | \psi_{v''} \rangle \times \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{v'+1} \dots \psi'_{N_v} | \psi_1 \dots \cancel{\psi_{v''}} \dots \psi_{v-1} \xi_v \psi_{v+1} \dots \psi_{N_v} \rangle \right| \\ & + (-1)^{v-1} \langle \psi'_{c'} | \xi_v \rangle \times \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{v'+1} \dots \psi'_{N_v} | \psi_1 \dots \psi_{v-1} \psi_{v+1} \dots \psi_{N_v} \rangle \right| \Bigg] \end{aligned} \tag{4.46}$$

Further refinement of this expression is achieved by splitting the sum into two parts based on the indices $v'' < v$ and $v'' > v$, accounting for the different scenarios in which $\psi_{v''}$ is removed:

$$\begin{aligned}\tilde{A}_{v',c'} = \sum_{v}(-1)^{v'-1}\Bigg[ &\sum_{v''<v}(-1)^{v-1}(-1)^{v''}\langle\psi'_{c'}|\psi_{v''}\rangle\times \\ &\det\left|\langle\psi'_1\dots\psi'_{v'-1}\psi'_{v'+1}\dots\psi'_{N_v}|\xi_v\psi_1\dots\cancel{\psi_{v''}}\dots\psi_{v-1}\psi_{v+1}\dots\psi_{N_v}\rangle\right| \\ &+\sum_{v''>v}(-1)^{v-1}(-1)^{v''-1}|\psi_{v''}\rangle\langle\psi'_{c'}|\psi_{v''}\rangle\times \\ &\det\left|\langle\psi'_1\dots\psi'_{v'-1}\psi'_{v'+1}\dots\psi'_{N_v}|\xi_v\psi_1\dots\cancel{\psi_{v''}}\dots\psi_{v-1}\psi_{v+1}\dots\psi_{N_v}\rangle\right| \\ +(-1)^{v-1}\langle\psi'_{c'}|\xi_v\rangle\times\det&\left|\langle\psi'_1\dots\psi'_{v'-1}\psi'_{v'+1}\dots\psi'_{N_v}|\psi_1\dots\psi_{v-1}\psi_{v+1}\dots\psi_{N_v}\rangle\right|\Bigg]\end{aligned} \tag{4.47}$$

The batches in the perturbed subspace are formulated analogously to those in the unperturbed subspace Eq.4.48 as:

$$|\tilde{\xi}'_{v'}\rangle = \sum_{c'}\tilde{A}_{v',c'}\,|\psi'_{c'}\rangle \tag{4.48}$$

Substituting Eq.4.47 into Eq.4.48

$$\begin{aligned}|\tilde{\xi}'_{v'}\rangle = \sum_{c'}|\psi'_{c'}\rangle\Bigg(\sum_{v}(-1)^{v'-1}\Bigg[&\sum_{v''<v}(-1)^{v-1}(-1)^{v''}\langle\psi'_{c'}|\psi_{v''}\rangle \\ &\times\det\left|\langle\psi'_1\dots\psi'_{v'-1}\psi'_{v'+1}\dots\psi'_{N_v}|\xi_v\psi_1\dots\cancel{\psi_{v''}}\dots\psi_{v-1}\psi_{v+1}\dots\psi_{N_v}\rangle\right| \\ &+\sum_{v''>v}(-1)^{v-1}(-1)^{v''-1}\langle\psi'_{c'}|\psi_{v''}\rangle \\ &\times\det\left|\langle\psi'_1\dots\psi'_{v'-1}\psi'_{v'+1}\dots\psi'_{N_v}|\xi_v\psi_1\dots\cancel{\psi_{v''}}\dots\psi_{v-1}\psi_{v+1}\dots\psi_{N_v}\rangle\right| \\ &+(-1)^{v-1}\langle\psi'_{c'}|\xi_v\rangle\det\left|\langle\psi'_1\dots\psi'_{v'-1}\psi'_{v'+1}\dots\psi'_{N_v}|\psi_1\dots\psi_{v-1}\psi_{v+1}\dots\psi_{N_v}\rangle\right|\Bigg]\Bigg)\end{aligned} \tag{4.49}$$

Recognizing the terms with the projection operator over the empty state manifold of

the perturbed KS Hamiltonian:

$$\hat{P}'_c = 1 - \sum_{v'=1,N_v} |\psi'_{v'}\rangle\langle\psi'_{v'}| = \sum_{c'} |\psi'_{c'}\rangle\langle\psi'_{c'}| \tag{4.50}$$

we obtain the un-normalized batches representation in the perturbed manifold

$$\begin{aligned}
|\tilde{\xi}'_v\rangle = \sum_v (-1)^{v'-1} \Bigg[ & \sum_{v''<v} (-1)^{v-1} (-1)^{v''} \hat{P}'_c |\psi_{v''}\rangle \\
& \times \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{v'+1} \dots \psi'_{N_v} | \xi_v \psi_1 \dots \not{\psi}_{v''} \dots \psi_{v-1} \psi_{v+1} \dots \psi_{N_v} \rangle \right| \\
& - \sum_{v''>v} (-1)^{v-1} (-1)^{v''} \hat{P}'_c |\psi_{v''}\rangle \\
& \times \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{v'+1} \dots \psi'_{N_v} | \xi_v \psi_1 \dots \not{\psi}_{v''} \dots \psi_{v-1} \psi_{v+1} \dots \psi_{N_v} \rangle \right| \\
& + (-1)^{v-1} \hat{P}'_c |\xi_v\rangle \det \left| \langle \psi'_1 \dots \psi'_{v'-1} \psi'_{v'+1} \dots \psi'_{N_v} | \psi_1 \dots \psi_{v-1} \psi_{v+1} \dots \psi_{N_v} \rangle \right| \Bigg]
\end{aligned} \tag{4.51}$$

The overlap of an excited state of the unperturbed Hamiltonian and the many-body ground state of the perturbed one Eq.4.40 is given by:

$$\begin{aligned}
g &= \langle \Psi'_0 | \Psi_S \rangle \\
&= \sum_v \det \langle \psi'_1 \dots \psi'_{N_v} | \psi_1 \dots \psi_{v-1} \xi_v \psi_{v+1} \dots \psi_{N_v} \rangle
\end{aligned} \tag{4.52}$$

Utilizing the Hellmann-Feynman theorem to obtain the forces of the excited state:

$$\begin{aligned}
\mathbf{F}_I^{exc} &= -\frac{\partial}{\partial \mathbf{R}_I} \langle \Psi_S | \hat{H}(\mathbf{R}_I)) | \Psi_S \rangle \\
&= -\langle \Psi_S | \frac{\partial}{\partial \mathbf{R}_I} \hat{H}(\mathbf{R}_I)) | \Psi_S \rangle
\end{aligned} \tag{4.53}$$

where, $\hat{H}$ is the Hamiltonian that describes the excited state, hence why the excitation energy is summed to the ground state energy as

$$\hat{H} = \hat{H}_S + E_0 |\Psi_0\rangle\langle\Psi_0| \tag{4.54}$$

The derivative of the excitonic Hamiltonian is obtained using finite differences as:

$$\frac{\partial}{\partial \mathbf{R}_I}\hat{H}_S(\mathbf{R}_I) = \frac{1}{2\Delta}(\hat{H}_S(\mathbf{R}_I + \Delta\hat{\mathbf{u}}) - \hat{H}_S(\mathbf{R}_I - \Delta\hat{\mathbf{u}})) \tag{4.55}$$

For the Hellman-Feynman theorem to be applicable the wavefunctions must be eignstates of the perturbed Hamiltonian as seen in App.B. Consequently, the excited state wavefunction within the perturbed manifold is utilized, as specified in Eq. 4.39.

$$\begin{aligned}\mathbf{F}_I^{exc} = -\frac{1}{2\Delta}\Big(&\langle\Psi_S|\hat{\mathcal{P}}(\mathbf{R}_I + \Delta\hat{\mathbf{u}})\hat{H}_S(\mathbf{R}_I + \Delta\hat{\mathbf{u}})\hat{\mathcal{P}}(\mathbf{R}_I + \Delta\hat{\mathbf{u}})|\Psi_S\rangle \\ &- \langle\Psi_S|\hat{\mathcal{P}}(\mathbf{R}_I - \Delta\hat{\mathbf{u}})\hat{H}_S(\mathbf{R}_I - \Delta\hat{\mathbf{u}})\hat{\mathcal{P}}(\mathbf{R}_I - \Delta\hat{\mathbf{u}})|\Psi_S\rangle\Big) \\ &+ \frac{1}{2\Delta}(g(\mathbf{R}_I + \Delta\hat{\mathbf{u}})^2 - g(\mathbf{R}_I - \Delta\hat{\mathbf{u}})^2)\mathbf{F}_I^0\end{aligned} \tag{4.56}$$

In this equation, $\mathbf{F}_I^0$ is the force calculated for the ground state. The squared overlap term is included as a correction, which accounts for the changes in the ground state due to the perturbation and their influences on electronic transitions and the resulting forces within the dynamics of the excited state. The forces obtained through Eq.4.56 are then used to optimize the excited state geometry.

## Constrained DFT

Constrained DFT allows the investigation of excited states by varying the occupation number $f_i$ in a localized state [79]. Investigations of the 1st excited state for a set of molecules using CDFT proved to be only slightly worse than linear-response time-dependent DFT (TDDFT) but better results compared to ΔSCF making it an appealing method in terms of computational cost and accuracy [89]. However, this approach is not easy to carry out in solids, since the concept of localized states whose occupation number can be varied is not well defined [36]. The electron density

$n(\mathbf{r})$ which is the key quantity in DFT can be written as:

$$n(\mathbf{r}) = \sum_i f_i \Psi_i^*(\mathbf{r}) \Psi_i(\mathbf{r}) \tag{4.57}$$

where, $f_i$ are the occupation numbers indicating how many electrons occupy each state i, it has the value of either 0 or 1.

The Kohn-Sham energy functional can be written as:

$$\begin{aligned} E[n] = \sum_{i=1}^{N} f_i \langle \Psi_i | T | \Psi_i \rangle + \frac{1}{2} + \int d\vec{r} d\vec{r'} \frac{n(\vec{r}) n(\vec{r'})}{|\vec{r} - \vec{r'}|} \\ + \int d\vec{r'} V_{ext}(\vec{r'}) n(\vec{r}) + \int d\vec{r} \varepsilon xc[n(\vec{r})] n(\vec{r}) \end{aligned} \tag{4.58}$$

The excited states using the energy functional Eq.4.58, are obtained by constraining a user-defined population of electrons, $f_i$, in the unoccupied level. To get the 1st excited state we impose $f_i = 1$

The force experienced by an atom due to its displacement can be obtained by taking

the derivative of the energy functional with respect to the atomic coordinate **R**:

$$\begin{aligned}
F =& -\frac{\partial E[n]}{\partial R} \\
=& -\sum_{i=1}^{N} f_i[\left\langle \frac{\partial \Psi_i}{\partial R} \right| T \left| \Psi_i \right\rangle + \left\langle \Psi_i \right| T \left| \frac{\partial \Psi_i}{\partial R} \right\rangle] \\
& - \int d\vec{r} \frac{n(\vec{r'})}{|\vec{r} - \vec{r'}|} \sum_{i}^{N} f_i[\frac{\partial \Psi_i^* \vec{r}}{\partial R} \Psi_i(\vec{r}) + \Psi_i^*(\vec{r}) \frac{\partial \Psi_i(\vec{r})}{\partial R}] \\
& - \int d\vec{r} \frac{\partial V_{ext}(\vec{r})}{\partial R} n(\vec{r}) \\
& - \int d\vec{r} V_{ext}(\vec{r}) \sum_{i}^{N} f_i[\frac{\partial \Psi_i^* \vec{r}}{\partial R} \Psi_i(\vec{r}) + \Psi_i^*(\vec{r}) \frac{\partial \Psi_i(\vec{r})}{\partial R}] \\
& - \int d\vec{r} \varepsilon xc[n](\vec{r}) \sum_{i}^{N} f_i[\frac{\partial \Psi_i^*(\vec{r})}{\partial R} \Psi_i(\vec{r}) + \Psi_i^*(\vec{r}) \frac{\partial \Psi_i(\vec{r})}{\partial R}] \\
& - \frac{\delta \varepsilon xc[n]}{\delta n(\vec{r})} n(\vec{r}) \sum_{i}^{N} f_i[\frac{\partial \Psi_i^*(\vec{r})}{\partial R} \Psi_i(\vec{r}) + \Psi_i^*(\vec{r}) \frac{\partial \Psi_i(\vec{r})}{\partial R}]
\end{aligned} \tag{4.59}$$

Note that in Eq.4.59 and unlike the standard HF for Kohn-Sham DFT we are taking the derivative of the wavefunctions. One thing one needs to pay attention to while evaluating the forces in Eq.4.59 is the wavefunctions that are being used. The wavefunctions derivative is obtained by finite difference method. We run two non-self consistent DFT calculations, for $R + \delta$ and $R - \delta$ with the changed population $f_i$. The resultant wavefunctions have to be rotated in a way that projection on the eigenfunctions matching the ones of the ground state is maximum.

If we have the following linear combination:

$$|\Psi_1\rangle = cos\theta \, |\phi_3\rangle + sin\theta \, |\phi_4\rangle \tag{4.60}$$

To find the angle of rotation, we want to have maximum projection over $|\phi_3\rangle$ and $|\phi_4\rangle$.

$$\langle \phi_3 | \Psi \rangle = cos\theta \, \langle \phi_3 | \phi_3 \rangle + sin\theta \, \langle \phi_3 | \phi_4 \rangle \tag{4.61}$$

$$\frac{d\langle\phi_3|\Psi\rangle}{d\theta} = 0 \tag{4.62}$$

The angle of rotation $\theta$ is given by:

$$\theta = \tan^{-1}\left(\frac{\langle\phi_3|\phi_4\rangle}{\langle\phi_3|\phi_3\rangle}\right) \tag{4.63}$$

## 4.3 Geometry optimization

A local quadratic approximation of the potential energy surface can be constructed by taking the Taylor expansion of the energy around a stationary point ($x_{k+1}$) as follows [90]:

$$E(x_{k+1}) - E(x_k) = g(x_k)^T s_k + \frac{1}{2} s_k^T H_k s_k \tag{4.64}$$

where, $g(X_k)$ is the gradient of the Energy function at $x_k$, defined as $g(x_k) = \nabla E(x_k) = -F(x_k)$, $H_k$ is the Hessian matrix at $x_k$, which represents the second-order partial derivatives of the Energy function. $s_k = x_{k+1} - x_k$ is the step taken from the current point $x_k$ to the next point $x_{k+1}$.

The quasi-Newton method alongside the Newton method are considered to be the most efficient and widely spread algorithms used for optimizing equilibrium geometries [18]. An approximation of the gradient $g(x)$ can be written as:

$$g(x_{k+1}) - g(x_k) = H_k s_k \tag{4.65}$$

The equilibrium geometry, we are seeking is given by the condition that the forces acting on individual nuclei vanish. This translates to the gradient being zero, i.e., $g(x_{k+1}) = 0$. Consequently, the displacement to the minima, also known as the

Newton-Raphson step is given by [18]:

$$s_k^{NR} = -H_k^{-1} g_k \tag{4.66}$$

In the quasi-Newton method the Hessian matrix is not evaluated at each iteration but rather, it gets systematically updated at each step of the optimization, which makes it more computationally appealing compared to the Newton-method. One of the most successful and widely used methods to update the Hessian matrix is the Broyden, Fletcher, Goldfarb, and Shanno (BFGS) [91, 92, 93, 94] update:

$$H_{k+1}^{-1} = H_k^{-1} + \left(1 + \frac{{\gamma_k}^T H_k^{-1} \gamma_k}{{s_k}^T \gamma_k}\right) \frac{s_k s_k^T}{{s_k}^T \gamma_k} - \left(\frac{s_k {\gamma_k}^T H_k^{-1} + H_k^{-1} \gamma_k s_k^T}{{s_k}^T \gamma_k}\right) \tag{4.67}$$

For minimization purposes the eigenvalues of the Hessian matrix should all be positive. At each step the position $x_k$ in the parameter space gets updated as follows:

$$x_{k+1} = x_k + T_k^L \frac{s_k^{NR}}{\|s_k^{NR}\|} \tag{4.68}$$

where $s_k^{NR}$ is given by Eq.4.66, $T_k^L$ is the trust radius which indicates the local region where the quadratic approximation is satisfactory.

Algorithm1 sketches quasi-Newton method with updating of the Hessian matrix.

**Algorithm 1** Quasi-Newton BFGS Method for Geometry Optimization

Initialize the geometry vector $x_0$ and the Hessian matrix $H_0$ (often set to the identity matrix)
Initialize the trust radius $T_0^L$
$k \leftarrow 0$
**while** not converged **do** ▷ Convergence based on a tolerance
    Compute the gradient $g_k = \nabla E(x_k)$
    **if** $\|g_k\|$ is below a predefined tolerance **then**
        **break**
    Compute the Newton-Raphson step $s_k^{NR} = -H_k^{-1} g_k$
    Update $x_{k+1} = x_k + T_k^L \frac{s_k^{NR}}{\|s_k^{NR}\|}$
    Compute $\gamma_k = g_{k+1} - g_k$
    Update $H_{k+1}^{-1}$ using the BFGS update formula
    Adjust the trust radius $T_{k+1}^L$ based on the reduction in energy
    $k \leftarrow k + 1$
**return** $x_k$ ▷ The optimized geometry vector

# Chapter Five

# From Theory to Practice: CO Molecule As A Test Run

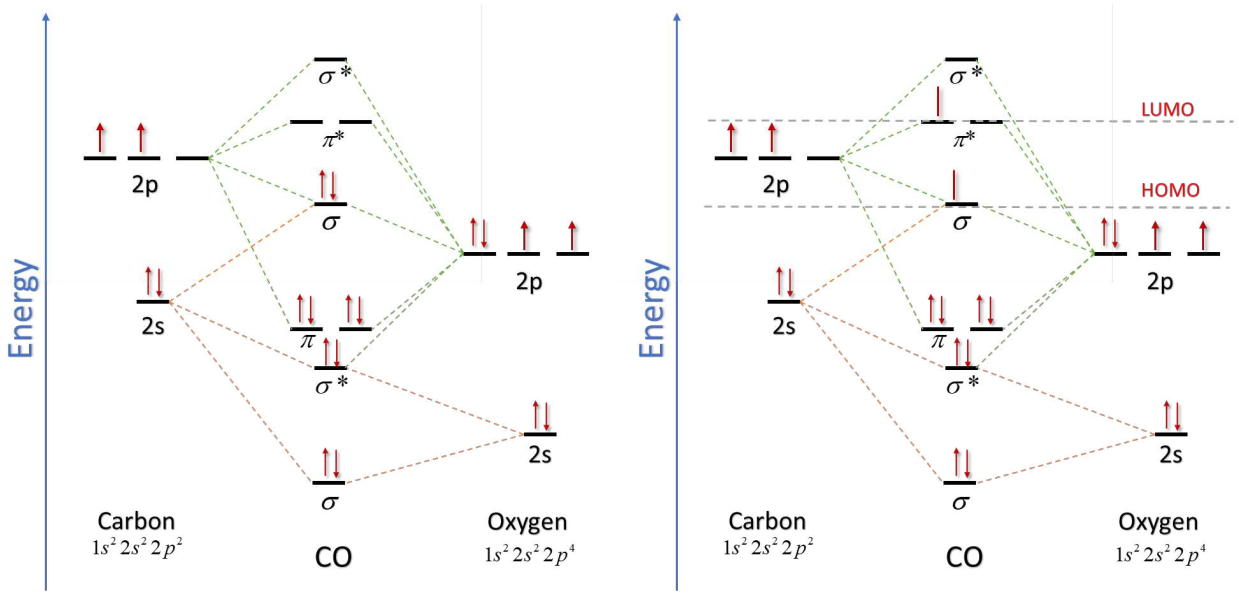
## 5.1 Computational Details

In order to simplify the calculations, we only consider the Gamma point (k=0) thus the wave function simplifies to the purely real part. All the programs employed in this study are part of the Quantum Espresso suite, a comprehensive collection of ab-initio codes designed for electronic structure computations [95]. Within this suite, the implementation of DFT methods is carried out using the 'pw' code, while the GW approximation and the BSE calculations are facilitated through the 'GWL' code [62, 63, 80].

We obtain the DFT wave functions and eigenvalues by solving the Kohn-Sham equations using the PBE exchange correlation functional [41] with a Norm-conserving pseudopotential [96]. We used a cubic simulation cell with a side length 20 bohr.

## 5.2 Results

Fig.5.1 illustrates the molecular orbital diagram of CO formed from the atomic orbitals of the carbon and Oxygen atoms. The ground state for the carbon monoxide molecule which possess a close-shell configuration.




**(a)** CO ground state electronic configuration. **(b)** CO excited state electronic configuration.

**Figure 5.1:** A sketch of energy level diagram showing the molecular orbital (MO) formation from the atomic orbitals of carbon and oxygen. The asterisk identifies anti-bonding orbitals.

In the context of molecular spectroscopy, and employing molecular term symbols, an electronic state of a diatomic molecule is typically denoted as:

$$^{2S+1}\Lambda^{(+/-)}$$

where S represents the total spin angular momentum of the electrons, and 2S+1 denotes the spin multiplicity. The symbol $\Lambda$ is the projection of the orbital angular momentum along the internuclear axis. The superscript $(+/-)$ indicates the reflection symmetry of the electronic wave function with respect to an arbitrary plane containing the internuclear axis, with $(+)$ for symmetric (gerade) and (-) for antisymmetric (ungerade) states.

For the CO molecule, the ground state configuration is described by the electronic term symbol $X^1\Sigma^+$. Here, $X$ indicates the ground state of the diatomic molecule, the superscript '1' denotes a singlet state with no unpaired electrons and thus a spin multiplicity of one. $\Sigma^+$ implies a **totally symmetric** state with zero orbital angular

momentum $\Lambda = 0$ and positive inversion symmetry.

The lowest lying electronic excited state results from the promotion of an electron from the highest occupied molecular orbital (HOMO), which is a bonding $\sigma$ orbital, to the lowest unoccupied molecular orbital (LUMO), which is a doubly degenerate anti-bonding $\pi^*$ orbital, as illustrated in Fig.5.1.b. This transition generates two unpaired electrons, one in the symmetric $\sigma$ orbital and one in the asymmetric $\pi^*$ orbital. The direct product of symmetries for these orbitals yields an overall $\Pi$ symmetry for the resulting molecular orbital. If the spins of the two unpaired electrons are parallel, a triplet state (S=1) results, denoted by $^3\Pi$. Conversely, if spins are antiparallel, the resulting state is a singlet state (S=0), denoted by $^1\Pi$.

According to electronic transition selection rules, transitions between states of the same spin multiplicity are allowed. Therefore, the allowed transition is $X^1\Sigma^+ \rightarrow A^1\Pi$ where A is the symbol for the first excited state. On the other hand, transitions to the triplet state $^3\Pi$ are spin-forbidden [97].

Fig.5.2 compares the KS energies obtained from DFT calculations and the quasi-particle energies obtained from GW approximation.

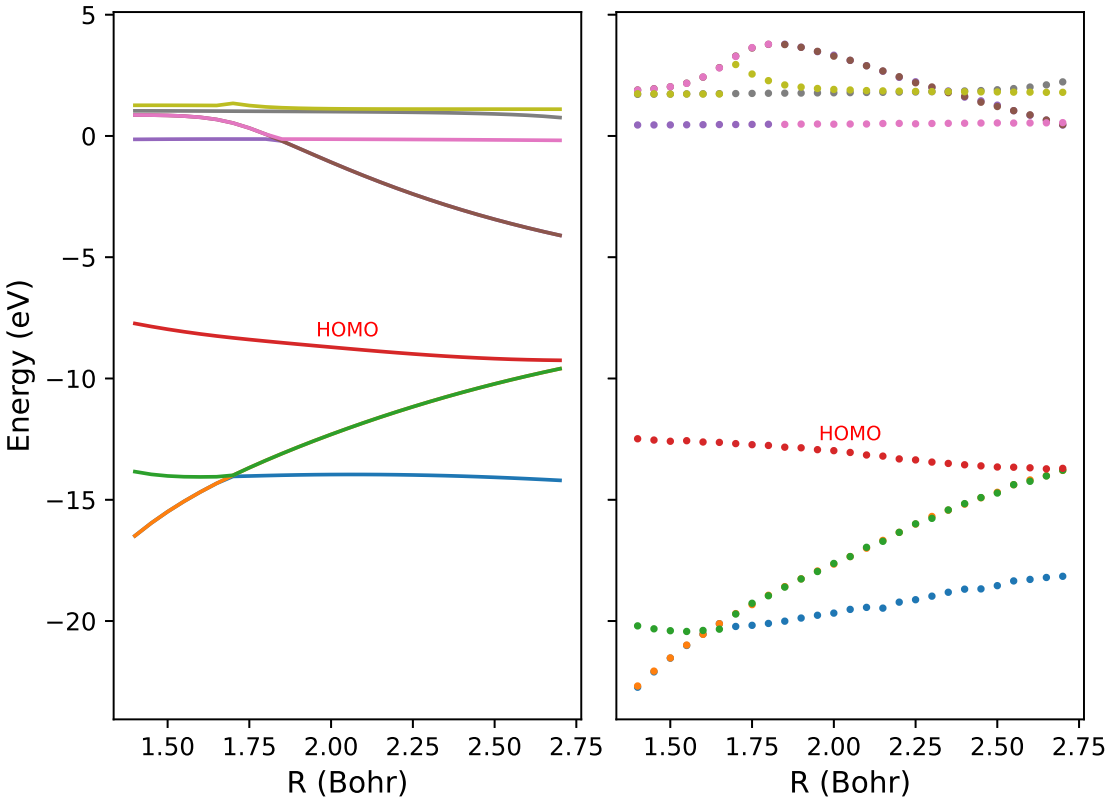



**Figure 5.2:** KS and quasiparticles energy obtained using DFT and GW calculations respectively. The deep lying energy levels were not considered in these plots.

The ionization potential for the ground state equilibrium geometry, which in our calculations was equal to 2.15 bohr, can be calculated from the KS-highest occupied molecular orbital (HOMO) as $IP = -\varepsilon_{HOMO}$. Here, the ionization potential is 8.71 eV which is too small, but that's to be expected since we aren't using an exact xc functional. The GW approximation value on the other hand was 13.6 eV which is in better agreement with the experimental value of 14.1 eV [98]. The discrepancy is due to the ambiguity of the vacuum level which results in shifting the quasi-particle energies.

In principle, for a diatomic molecule like CO with only one degree of freedom obtaining the optimized geometry of the $1^{st}$ excited state can be done by studying the potential energy surface of the excited state total energy Eq.4.54 and from the

minima of the surface corresponds to the optimized bond length. Fig.5.3 shows a number of excited state levels obtained using GW-BSE approach within the GWL code as a function of the bond length.

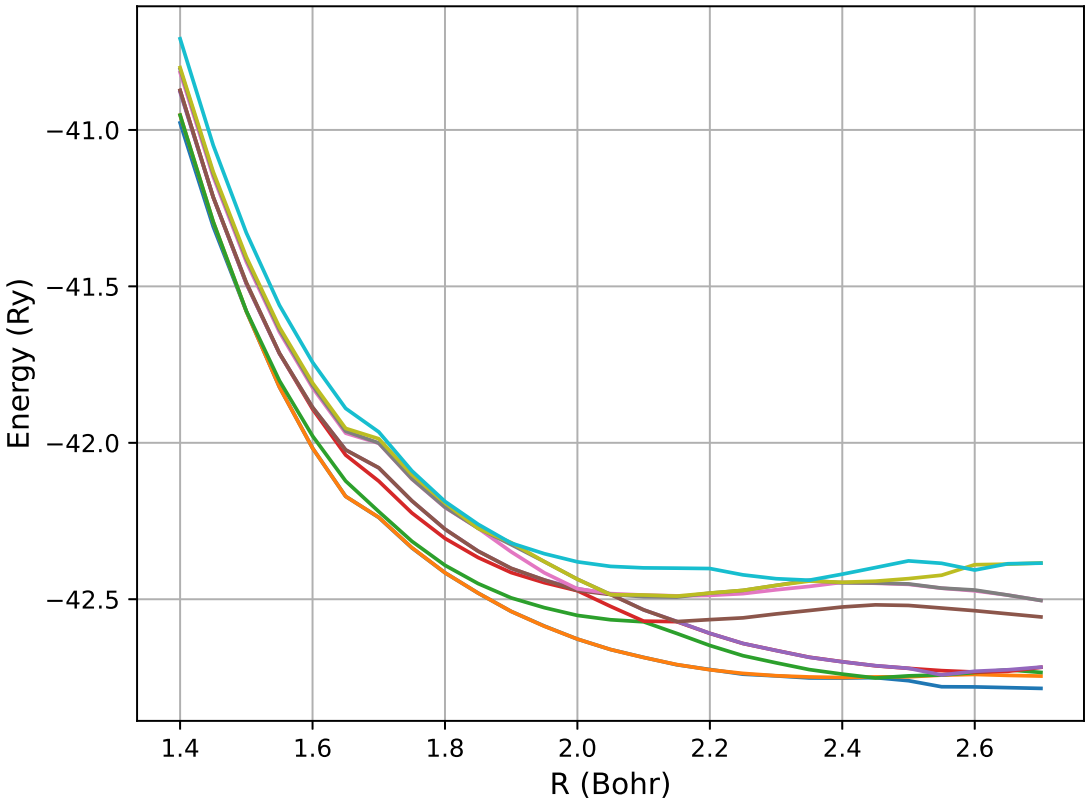



**Figure 5.3:** Excited state total energies for CO molecule as a function of bond length calculated with GW-BSE.

While investigating the potential energy surfaces of the excited states one need to pay extra attention to keep track of the correct state, as we could have cases of energy crossing as illustrated in Fig.5.3.

We also investigated the 1st excited state of CO, by promoting an electron from the HOMO to LUMO, we ran non-self consistent calculations, unlike CDFT where you have to iterate to self-consistency, Fig.5.4 shows the 1st excited energy levels as a function of the bond length for the $A^1\Pi$ excited state.

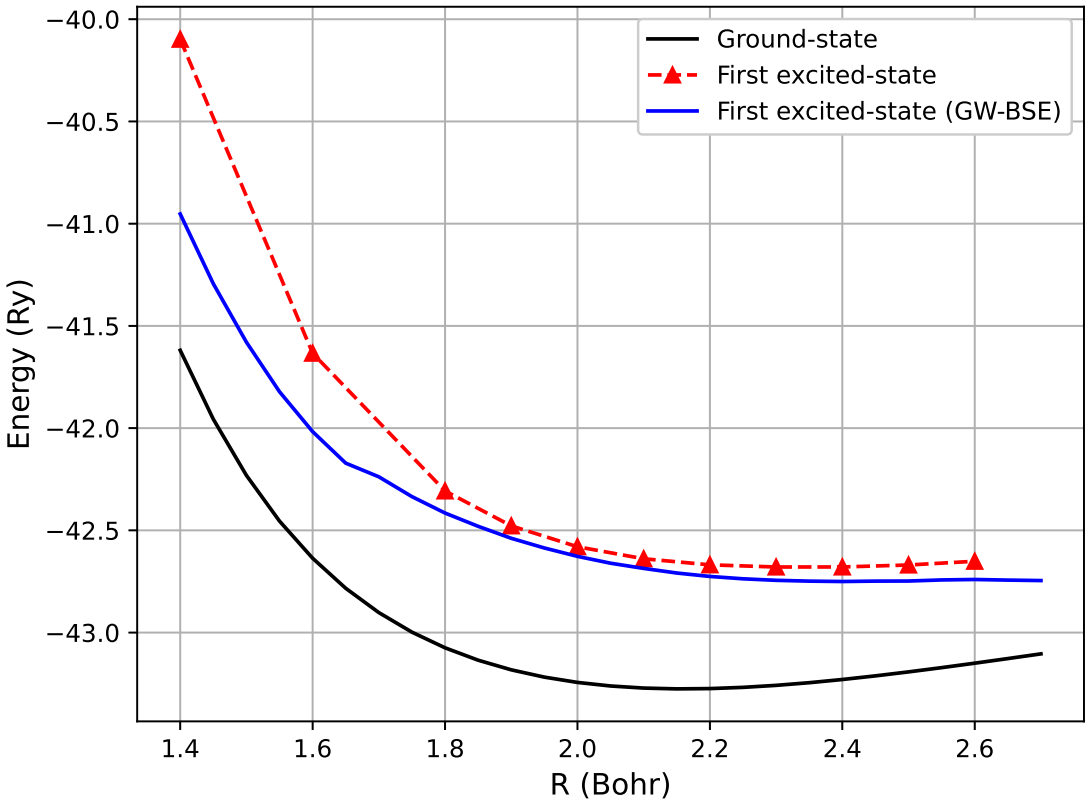



**Figure 5.4:** Ground state $^1\Sigma^+$, $A^1\Pi$ excited state energies of CO as a function of the bond length. The triangle are the energies obtained from running non-self consistent constrained DFT

As we mentioned previously the optimized bond length of CO can be obtained right away once the potential energy surface is available. However, that is impractical for molecules with more degrees of freedom, hence why it is interesting to evaluate the gradient of energy and use it for structural optimization by utilizing one of the various optimization techniques, e.g BFGS that we discussed briefly in Sec.4.3.

Fig.5.5 shows the forces acting on the C atom for various bound lengths for the three lowest excited levels obtained Using Eq.4.56.

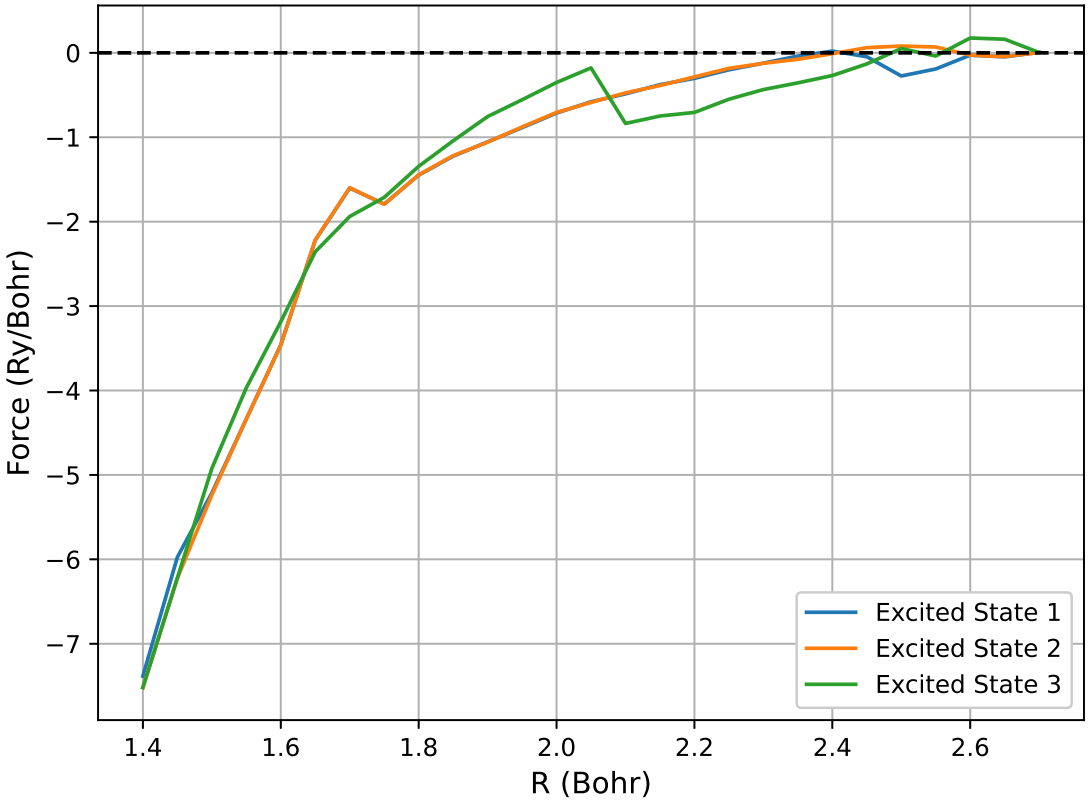



**Figure 5.5:** The force on Carbon atom in CO as a function of the bond length R.

It is interesting to point out the behaviour of the curve for bond lengths (2.4 - 2.6) bohr in Fig.5.8. It is due to the energy level crossing as illustrated by Fig.5.3. We have the equilibrium geometry at 2.4 bohr, the dip in the force curve is due to the fact that the forces are due to other levels. We have another equilibrium geometry at 2.6 bohr which corresponds to another energy level. Also between 1.6-1.8 bohr, which can be traced back to the level crossing as illustrated by Fig.5.3.

We can obtain the forces within constrained DFT, by using Hellmann-Feynman theorem as illustrated in Sec.4.2. In this work we used a non-self consistent calculations. Unlike standard HF theorem where the derivative of the wavefunction

vanishes when using the exact eigenfunctions of the Hamiltonian. Here, it's not the case, the derivative of the wave functions plays a significant role in determining the forces as illustrated in Fig.5.6. The derivative of the wavefunctions was obtained using finite differences, which makes this approach quite impractical.

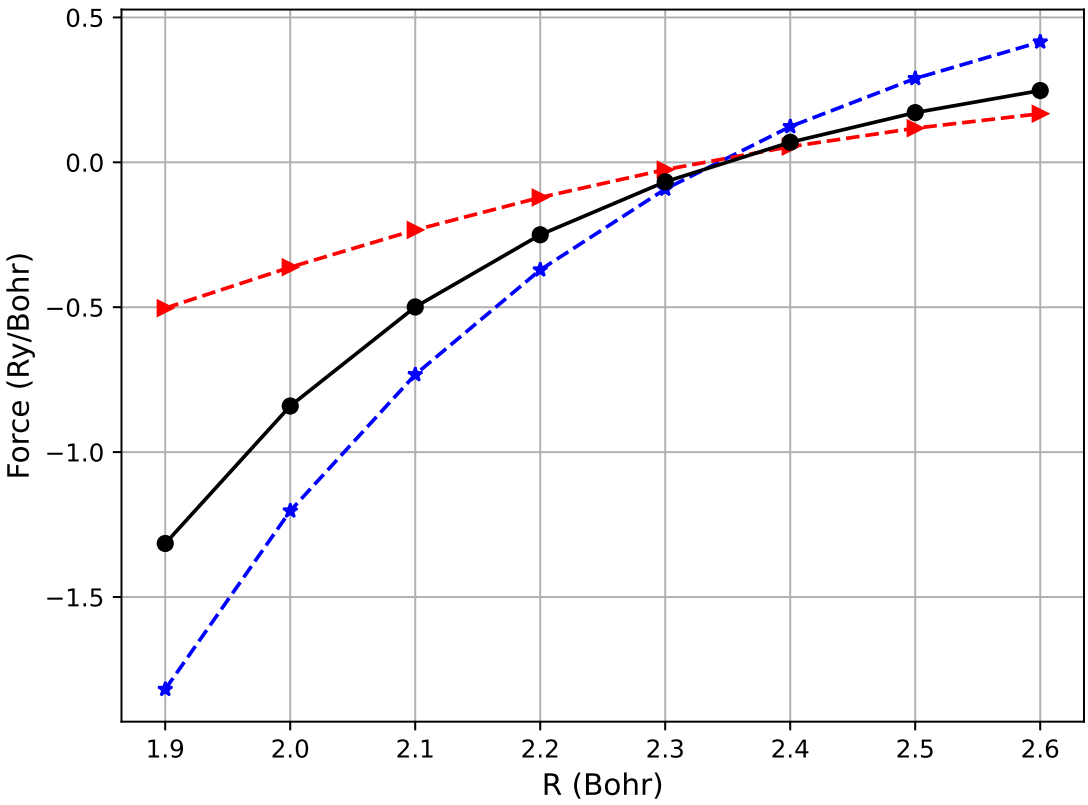



**Figure 5.6:** The force within HF-CDFT on Carbon atom in CO as a function of the bond length R. The blue stars are the HF-Forces without the wavefunctions corrections, the red triangles are the wavefunctions correction to the HF theorem. The black dots are the total HF-forces with wavefunction correction taken into account. We used a non-self consistent CDFT for comparison purposes

Fig.5.7 shows the forces acting on the C atom obtained using GW-BSE and the ones from running a non-self consistent constrained DFT. Both methods predict an an optimized bond length which are comparable to the experimental value. The results obtained are summarized in Table 5.1 which also compares them to various results reported in literature.

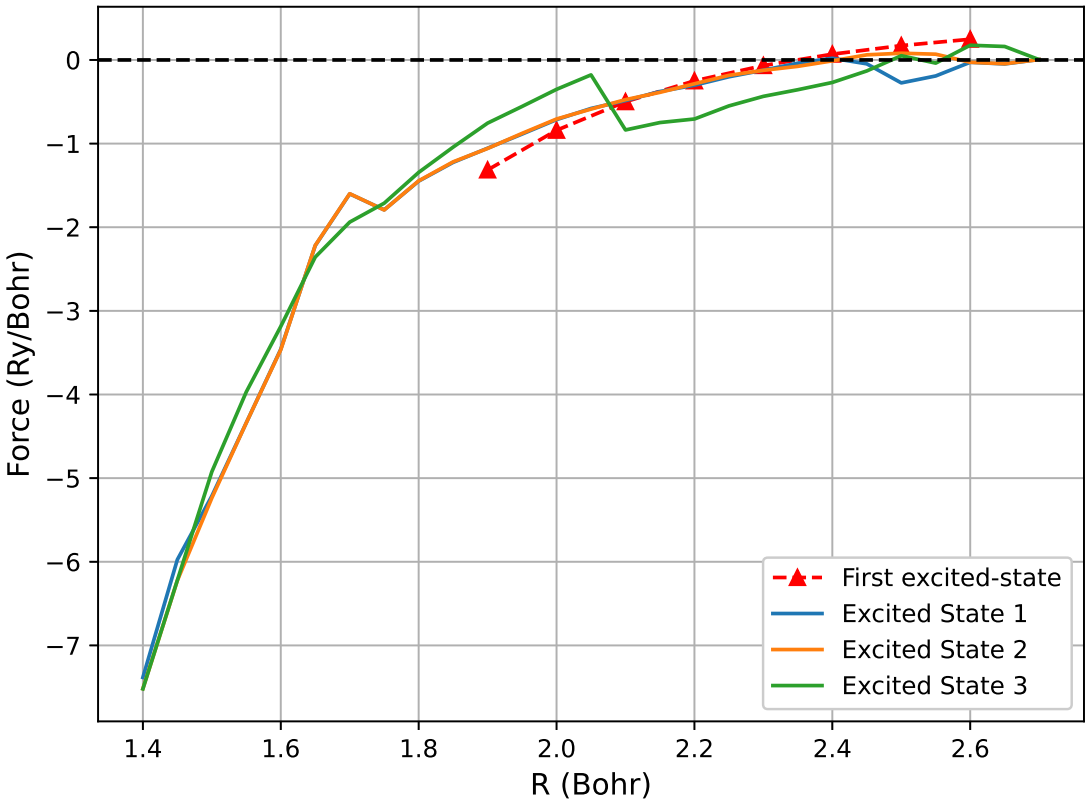



**Figure 5.7:** Forces on the C atom for the CO molecule obtained using GW-BSE and the non-self consistent constrained DFT

**Table 5.1:** Bond length for CO molecule in the $A^1\Pi$ excited state obtained from different studies in units of bohr

| GW-BSE | TDLDA | CDFT | CIS | Expt. |
|---|---|---|---|---|
| 1.25[1] | 1.25[1] | 1.21[2] | 1.213[4] | 1.235[a] |
| 1.26[2] | | | | |
| 1.249[3] | | | | |
| 1.27[*] | | | | |

[*] This work
[1] Ref.[99]
[2] Ref.[84]
[3] Ref.[100] for G0W0-BSE within TDA
[4] Ref.[101] for pVTZ+ basis
[a] Experimental data Ref.[102]

Ref.[100] offers a more comprehensive comparison of various computational methods

for the optimized bond length for the CO molecule in the $A^1\Pi$ excited state.

In order to determine how accurate the implementation of the approach presented in Sec.4.2 was, we compare the result we obtained, to the ones calculated by taking the two-points finite differences of the excited energies shown in Fig.5.3 as:

$$F = -\frac{\partial E_n}{\partial R} = \frac{E_n(R+\delta) - E_n(R-\delta)}{2\delta} \tag{5.1}$$

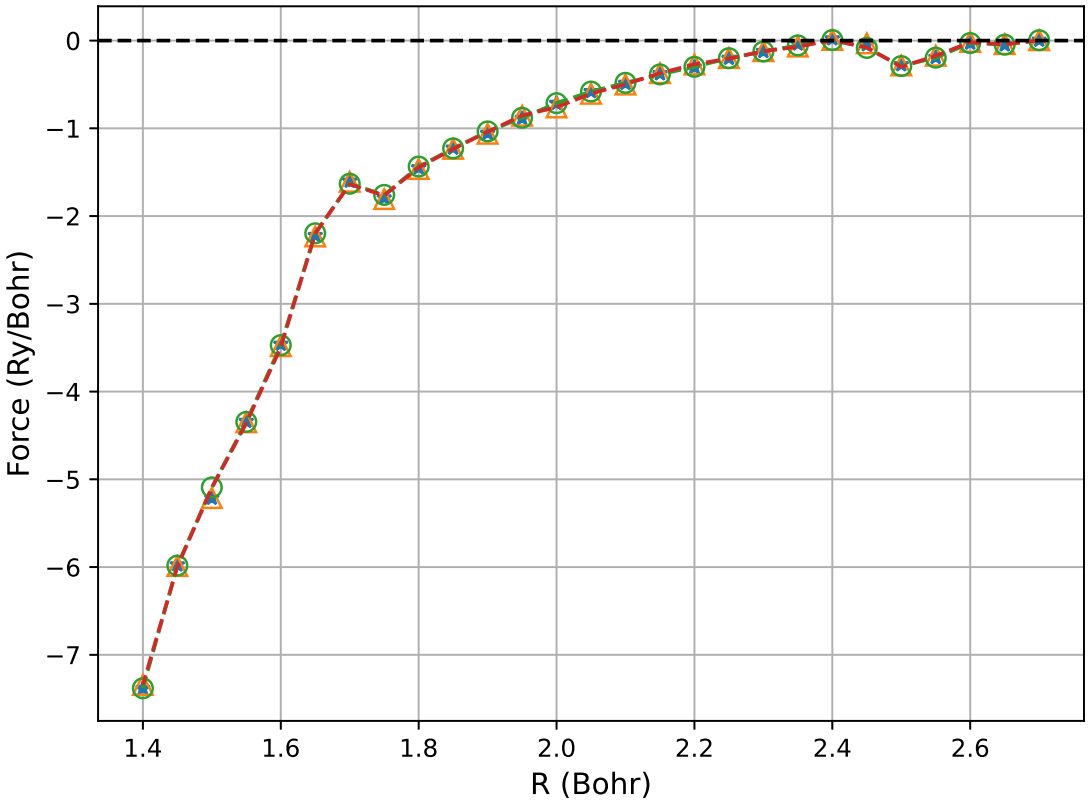



**Figure 5.8:** The first excited state forces on Carbon atom in CO as a function of the bond length R. The triangles and stars are the forces on the C atom and - force on the O atom respectively calculated using Eq.4.56 and the finite difference of the total energies

Fig.5.8 shows the forces acting on both C and O obtained using Eq.4.56 and the ones from Eq.5.1. The two results are in perfect agreement. Forces obtained by taking the finite differences Eq.5.1 are often referred to as "exact" forces in literature.

Fig.5.8 shows the forces acting on both C and O obtained using Eq.4.56 and the "exact" ones as a function of varying bond length.

## 5.3 Conclusions

In this work we have proposed a way to obtain the gradient of the excited state energy within the TDA-BSE approach Eq.4.56, which yields gradient of energies with perfect consistency with the finite differences way of obtaining the gradient. CO molecule made the perfect test subject due to the fact that it has only one degree of freedom allowing for careful study, as well as the available results within literature, one being the pioneering work by Ismail [84] which allows us to cross-check with other references. The pros of using the CO could also be thought of as it cons. Geometry optimization of a molecule with only one degree of freedom is not that compelling. Hence why we want to run the calculations for the lowest singlet excited states of the formaldehyde molecule $CH_2O$. $CH_2O$ makes an attractive test subject as it plays an important role in atmospheric chemistry [103]. It is the carbonyl compound with the most abundancy and is used to determine if the air quality is poor or not [104]. Further more it's geometry makes it an interesting case for study. In its ground state Formaldehyde poses the $C_{2\nu}$ symmetry [105]. Many studies reported that it undergoes significant geometrical changes when excited to $\pi^*$ state [106]. Dash et.al [107] have investigated the low-lying $n \rightarrow \pi^*$ excitation of Formaldehyde. They reported that for the optimal excited state structure, the symmetry is reduced from $C_{2\nu}$ to $C_s$. They also reported that the excited state geometry depends significantly on the level of theory employed to do the calculations. The geometry optimization will be carried out using the quasi-Newton method of Broyden, Fletcher, Goldfarb, and Shanno (BFGS) [91, 92, 93, 94] method implemented within Scipy-Python library [108, 109] with the gradient of energy calculated using Eq.4.56. Later on we plan

to carry the calculations for more diatomic molecules such as $H_2, C_2$ and $N_2$. Also for ammonia $NH_3$, a molecule with 6 degrees of freedom which makes the geometry optimization a bit more challenging.

# Appendix

# A Green's function identity

The Green's function is defined as follows:

$$\int d3G(1,3)G^{-1}(3,2) = \delta(1,2) \tag{A1}$$

Taking the functional derivative with respect to U(4) gives us:

$$\int d3\frac{\delta G(1,3)}{\delta U(4)}G^{-1}(3,2) + \int d3G(1,3)\frac{\delta G^{-1}(3,2)}{\delta U(4)} = 0 \tag{A2}$$

Multiply both sides by G(2,5)

$$\int d(23)\frac{\delta G(1,3)}{\delta U(4)}G^{-1}(3,2)G(2,5) + \int d(23)G(1,3)\frac{\delta G^{-1}(3,2)}{\delta U(4)}G(2,5) = 0 \tag{A3}$$

Making use of the inverse definition Eq.A1 and integrating over the variable (2) we get:

$$\frac{\delta G(1,5)}{U(4)} = -\int d(23)G(1,3)\frac{\delta G^{-1}(3,2)}{\delta U(4)}G(6,2) \tag{A4}$$

# B Forces within Hellmann-Feynman Theorem

The Hellmann-Feynman Theorem provides a fundamental link between quantum mechanical forces and the derivatives of energy with respect to external parameters [110]. To derive a general expression for this theorem, we start with the eigenvalue equation:

$$\hat{H}(\lambda)\Psi_n = E_n\Psi_n \tag{B1}$$

Taking the derivative with respect to the parameter $\lambda$:

$$\hat{H}(\lambda)\left|\frac{\partial\Psi_n}{\partial\lambda}\right\rangle + \frac{\partial\hat{H}}{\partial\lambda}\left|\Psi_n\right\rangle = E_n\left|\frac{\partial\Psi_n}{\partial\lambda}\right\rangle + \frac{\partial E_n}{\partial\lambda}\left|\Psi_n\right\rangle \tag{B2}$$

Multiplying both sides with $\langle\Psi_m|$ an eigenstate of $\hat{H}$ we obtain:

$$\langle\Psi_m|\,\hat{H}(\lambda)\left|\frac{\partial\Psi_n}{\partial\lambda}\right\rangle + \langle\Psi_m|\,\frac{\partial\hat{H}}{\partial\lambda}\,|\Psi_n\rangle = E_n\left\langle\Psi_m\middle|\frac{\partial\Psi_n}{\partial\lambda}\right\rangle + \frac{\partial E_n}{\partial\lambda}\,\langle\Psi_m|\Psi_n\rangle \tag{B3}$$

rearranging the equation:

$$\langle\Psi_m|\,\frac{\partial\hat{H}(\lambda)}{\partial\lambda}\,|\Psi_n\rangle = (E_n - E_m)\left\langle\Psi_m\middle|\frac{\partial\Psi_n}{\partial\lambda}\right\rangle + \frac{\partial E_n}{\partial\lambda}\delta_{mn} \tag{B4}$$

For $m = n$ we get the diagonal form of the Hellmann-Feynman theorem [111]:

$$\langle\Psi_n|\,\frac{\partial\hat{H}(\lambda)}{\partial\lambda}\,|\Psi_n\rangle = \frac{\partial E_n}{\partial\lambda} \tag{B5}$$

The Forces $F$ on atom i can be expressed as:

$$\vec{F}_{n,i} = -\frac{\partial E_n}{\partial\mathbf{R}_i} = -\,\langle\Psi_n|\,\frac{\partial\hat{H}}{\partial\mathbf{R}_i}\,|\Psi_n\rangle \tag{B6}$$

where $\mathbf{R}_i$ represents the position of atom i, the force is determined by the derivative of the Hamiltonian with respect to this position.

# Bibliography


[1] Feliciano Giustino. ***Materials modelling using density functional theory: properties and predictions***. Oxford University Press, 2014.

[2] Giovanni Onida, Lucia Reining, and Angel Rubio. "Electronic excitations: density-functional versus many-body Green's-function approaches." In: *Reviews of modern physics* 74.2 (2002), p. 601.

[3] Dorothea Golze, Marc Dvorak, and Patrick Rinke. "The GW compendium: A practical guide to theoretical photoemission spectroscopy." In: *Frontiers in chemistry* 7 (2019), p. 377.

[4] Mark van Schilfgaarde, Takao Kotani, and Sergey Faleev. "Quasiparticle self-consistent g w theory." In: *Physical review letters* 96.22 (2006), p. 226402.

[5] Stephan Hüfner. *Photoelectron spectroscopy: principles and applications*. Springer Science & Business Media, 2013.

[6] Xavier Blase, Ivan Duchemin, and Denis Jacquemin. "The Bethe–Salpeter equation in chemistry: relations with TD-DFT, applications and challenges." In: *Chemical Society Reviews* 47.3 (2018), pp. 1022–1043.

[7] Guglielmo Lanzani. *The photophysics behind photovoltaics and photonics*. John Wiley & Sons, 2012.

[8] Rajesh Kancherla et al. "Visible light-induced excited-state transition-metal catalysis." In: *Trends in Chemistry* 1.5 (2019), pp. 510–523.

[9] Volkhard Helms. *Principles of computational cell biology: from protein complexes to cellular networks*. John Wiley & Sons, 2018.

[10] Gianluca Tirimbò and Björn Baumeier. "Ab initio modeling of excitons: from perfect crystals to biomaterials." In: *Advances in Physics: X* 6.1 (2021), p. 1912638.

[11] Ching W Tang and Steven A VanSlyke. "Organic electroluminescent diodes." In: *Applied physics letters* 51.12 (1987), pp. 913–915.

[12] Björn Baumeier et al. "Challenges for in silico design of organic semiconductors." In: *Journal of Materials Chemistry* 22.22 (2012), pp. 10971–10976.

[13] Michael Rohlfing and Steven G Louie. "Electron-hole excitations and optical spectra from first principles." In: *Physical Review B* 62.8 (2000), p. 4927.

[14] Catalin D Spataru et al. "Excitonic effects and optical spectra of single-walled carbon nanotubes." In: *Physical Review Letters* 92.7 (2004), p. 077402.

[15] Xia Leng et al. "GW method and Bethe–Salpeter equation for calculating electronic excitations." In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 6.5 (2016), pp. 532–550.

[16] Bastian Schaefer et al. "Stabilized quasi-Newton optimization of noisy potential energy surfaces." In: *The Journal of chemical physics* 142.3 (2015).

[17] Benjamin Kaduk, Takashi Tsuchimochi, and Troy Van Voorhis. "Analytic energy gradients for constrained DFT-configuration interaction." In: *The Journal of Chemical Physics* 140.18 (2014).

[18] H Bernhard Schlegel. "Geometry optimization." In: *Wiley Interdisciplinary Reviews: Computational Molecular Science* 1.5 (2011), pp. 790–809.

[19] Philip W Anderson. "More Is Different: Broken symmetry and the nature of the hierarchical structure of science." In: *Science* 177.4047 (1972), pp. 393–396.

[20] Richard D Mattuck. *A guide to Feynman diagrams in the many-body problem*. Courier Corporation, 1992.

[21] Michael Springborg. *Methods of electronic-structure calculations: from molecules to solids*. 2000.

[22] Erwin Schrödinger. “Quantisierung als eigenwertproblem.” In: ***Annalen der physik*** 385.13 (1926), pp. 437–490.

[23] Robert B Laughlin and David Pines. “The theory of everything.” In: *Proceedings of the national academy of sciences* 97.1 (2000), pp. 28–31.

[24] Richard P Feynman et al. “Simulating physics with computers.” In: *Int. j. Theor. phys* 21.6/7 (2018).

[25] M Born and R Oppenheimer. “Zur Quantentheorie der Molekeln Annalen der Physik, v. 84.” In: (1927).

[26] John C Slater. “The ferromagnetism of nickel.” In: ***Physical Review*** 49.7 (1936), p. 537.

[27] Vladimir Fock. “Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems.” In: ***Zeitschrift für Physik*** 61.1-2 (1930), pp. 126–148.

[28] N.W. Ashcroft and N.D. Mermin. ***Solid State Physics***. Philadelphia: Saunders College, 1976.

[29] Charles Kittel and Ching-yao Fong. ***Quantum theory of solids***. Vol. 5. Wiley New York, 1963.

[30] Richard Van Noorden, Brendan Maher, and Regina Nuzzo. “The top 100 papers.” In: *Nature News* 514.7524 (2014), p. 550.

[31] Nathan Argaman and Guy Makov. “Density functional theory: An introduction.” In: ***American Journal of Physics*** 68.1 (2000), pp. 69–79.

[32] Walter Kohn. “Nobel Lecture: Electronic structure of matter—wave functions and density functionals.” In: ***Reviews of Modern Physics*** 71.5 (1999), p. 1253.

[33] Llewellyn H Thomas. "The calculation of atomic fields." In: *Mathematical Proceedings of the Cambridge Philosophical Society*. Vol. 23. 5. Cambridge University Press. 1927, pp. 542–548.

[34] Enrico Fermi. "A statistical method for determining some properties of the atom and their application to the theory of the periodic system of elements." In: ***Journal of Physics*** 48.1-2 (1928), pp. 73–79.

[35] Pierre Hohenberg and Walter Kohn. "Inhomogeneous electron gas." In: ***Physical review*** 136.3B (1964), B864.

[36] Richard M Martin. ***Electronic structure: basic theory and practical methods***. Cambridge university press, 2004.

[37] Walter Kohn and Lu Jeu Sham. "Self-consistent equations including exchange and correlation effects." In: ***Physical review*** 140.4A (1965), A1133.

[38] NM Harrison. "**An introduction to density functional theory**." In: (2003).

[39] Karlheinz Schwarz. "Computation of materials properties at the atomic scale." In: *Selected Topics in Applications of Quantum Mechanics*. InTech, 2015.

[40] Frank Herman, John P Van Dyke, and Irene B Ortenburger. "Improved statistical exchange approximation for inhomogeneous many-electron systems." In: ***Physical Review Letters*** 22.16 (1969), p. 807.

[41] John P Perdew, Kieron Burke, and Matthias Ernzerhof. "Generalized gradient approximation made simple." In: ***Physical review letters*** 77.18 (1996), p. 3865.

[42] Stephan Mohr et al. "Accurate and efficient linear scaling DFT calculations with universal applicability." In: *Physical Chemistry Chemical Physics* 17.47 (2015), pp. 31360–31370.

[43] RD King-Smith, MC Payne, and JS Lin. "Real-space implementation of nonlocal pseudopotentials for first-principles total-energy calculations." In: *Physical Review B* 44.23 (1991), p. 13063.

[44] Pedro Borlido et al. "Exchange-correlation functionals for band gaps of solids: benchmark, reparametrization and machine learning." In: *npj Computational Materials* 6.1 (2020), p. 96.

[45] James F Janak. "Proof that E n i= $\varepsilon$ in density-functional theory." In: *Physical Review B* 18.12 (1978), p. 7165.

[46] John P Perdew and Mel Levy. "Physical content of the exact Kohn-Sham orbital energies: band gaps and derivative discontinuities." In: *Physical Review Letters* 51.20 (1983), p. 1884.

[47] Richard M Martin, Lucia Reining, and David M Ceperley. *Interacting electrons*. Cambridge University Press, 2016.

[48] Alexander L Fetter and John Dirk Walecka. *Quantum theory of many-particle systems*. Courier Corporation, 2012.

[49] Adrian Stan et al. "Unphysical and physical solutions in many-body theories: from weak to strong correlation." In: *New Journal of Physics* 17.9 (2015), p. 093045.

[50] Viktor Mikhailovich Galitskii and Arkadii Beinusovich Migdal. "Application of quantum field theory methods to the many body problem." In: *Sov. Phys. JETP* 7.96 (1958), p. 18.

[51] Harry Lehmann. "Üon properties of propagation functions and renormalization constants of quantized fields." In: *Il Nuovo Cimento (1943-1954)* 11 (1954), pp. 342–357.

[52] Ferdi Aryasetiawan and Olle Gunnarsson. "The GW method." In: *Reports on Progress in Physics* 61.3 (1998), p. 237.

[53] Paul C Martin and Julian Schwinger. "Theory of many-particle systems. I." In: *Physical Review* 115.6 (1959), p. 1342.

[54] Freeman J Dyson. "The radiation theories of Tomonaga, Schwinger, and Feynman." In: *Physical Review* 75.3 (1949), p. 486.

[55] Lars Hedin. "New method for calculating the one-particle Green's function with application to the electron-gas problem." In: *Physical Review* 139.3A (1965), A796.

[56] Lars Hedin and Stig Lundqvist. "Effects of electron-electron and electron-phonon interactions on the one-electron states of solids." In: *Solid state physics*. Vol. 23. Elsevier, 1970, pp. 1–181.

[57] M Palummo et al. "Screening models and simplified GW approaches: Si & GaN as test cases." In: *Solid state communications* 95.6 (1995), pp. 393–398.

[58] PA Sterne and JC Inkson. "Exchange-correlation potential in semiconductors and insulators." In: *Journal of Physics C: Solid State Physics* 17.9 (1984), p. 1497.

[59] C Delerue, M Lannoo, and G Allan. "Calculations of the electron-energy-loss spectra of silicon nanostructures and porous silicon." In: *Physical Review B* 56.23 (1997), p. 15306.

[60] J Furthmüller et al. "GW self-energy calculations for systems with huge supercells." In: *Physical Review B* 66.4 (2002), p. 045110.

[61] Mark S Hybertsen and Steven G Louie. "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies." In: *Physical Review B* 34.8 (1986), p. 5390.

[62] P Umari, Geoffrey Stenuit, and Stefano Baroni. "Optimal representation of the polarization propagator for large-scale G W calculations." In: *Physical Review B* 79.20 (2009), p. 201104.

[63] Paolo Umari, Geoffrey Stenuit, and Stefano Baroni. "GW quasiparticle spectra from occupied states only." In: *Physical Review B* 81.11 (2010), p. 115104.

[64] R Sternheimer. "On nuclear quadrupole moments." In: *Physical Review* 84.2 (1951), p. 244.

[65] Xavier Blase, Claudio Attaccalite, and Valerio Olevano. "First-principles GW calculations for fullerenes, porphyrins, phtalocyanine, and other molecules of interest for organic photovoltaic applications." In: *Physical Review B* 83.11 (2011), p. 115103.

[66] Olga Dolgounitcheva et al. "Accurate ionization potentials and electron affinities of acceptor molecules IV: electron-propagator methods." In: *Journal of chemical theory and computation* 12.2 (2016), pp. 627–637.

[67] Giancarlo Strinati. "Application of the Green's functions method to the study of the optical properties of semiconductors." In: *La Rivista del Nuovo Cimento (1978-1999)* 11.12 (1988), pp. 1–86.

[68] Giovanna Lani, Pina Romaniello, and Lucia Reining. "Approximations for many-body Green's functions: insights from the fundamental equations." In: *New Journal of Physics* 14.1 (2012), p. 013056.

[69] Edwin E Salpeter and Hans Albrecht Bethe. "A relativistic equation for bound-state problems." In: *Physical Review* 84.6 (1951), p. 1232.

[70] Michael Rohlfing and Steven G Louie. "Electron-hole excitations in semiconductors and insulators." In: *Physical review letters* 81.11 (1998), p. 2312.

[71] Xavier Blase et al. "The Bethe–Salpeter equation formalism: From physics to chemistry." In: *The Journal of Physical Chemistry Letters* 11.17 (2020), pp. 7371–7382.

[72] W Hanke and LJ Sham. "Many-particle effects in the optical spectrum of a semiconductor." In: *Physical Review B* 21.10 (1980), p. 4656.

[73] S Albrecht et al. "Excitonic effects in the optical properties." In: *physica status solidi (a)* 170.2 (1998), pp. 189–197.

[74] Paul Boulanger et al. "Fast and accurate electronic excitations in cyanines with the many-body Bethe–Salpeter approach." In: *Journal of Chemical Theory and Computation* 10.3 (2014), pp. 1212–1218.

[75] Jeffrey C Grossman et al. "High accuracy many-body calculational approaches for excitations in molecules." In: *Physical Review Letters* 86.3 (2001), p. 472.

[76] Giovanni Onida et al. "Ab initio calculations of the quasiparticle and absorption spectra of clusters: the sodium tetramer." In: *Physical review letters* 75.5 (1995), p. 818.

[77] Tobias Sander, Emanuele Maggio, and Georg Kresse. "Beyond the Tamm-Dancoff approximation for extended systems using exact diagonalization." In: *Physical Review B* 92.4 (2015), p. 045209.

[78] Andrea Marini and Rodolfo Del Sole. "Dynamical excitonic effects in metals and semiconductors." In: *Physical review letters* 91.17 (2003), p. 176402.

[79] Friedhelm Bechstedt. *Many-body approach to electronic excitations*. Springer, 2016.

[80] Margherita Marsili et al. "Large-scale G W-BSE calculations with N 3 scaling: Excitonic effects in dye-sensitized solar cells." In: *Physical Review B* 95.7 (2017), p. 075415.

[81] Nicola Marzari and David Vanderbilt. "Maximally localized generalized Wannier functions for composite energy bands." In: *Physical review B* 56.20 (1997), p. 12847.

[82] Brent Walker et al. “Efficient approach to time-dependent density-functional perturbation theory for optical spectroscopy.” In: *Physical review letters* 96.11 (2006), p. 113001.

[83] Dario Rocca, Deyu Lu, and Giulia Galli. “Ab initio calculations of optical absorption spectra: Solution of the Bethe–Salpeter equation within density matrix perturbation theory.” In: *The Journal of chemical physics* 133.16 (2010).

[84] Sohrab Ismail-Beigi and Steven G Louie. “Excited-state forces within a first-principles Green’s function formalism.” In: *Physical review letters* 90.7 (2003), p. 076401.

[85] Pierre-François Loos et al. “Pros and cons of the Bethe–Salpeter formalism for ground-state energies.” In: *The journal of physical chemistry letters* 11.9 (2020), pp. 3536–3545.

[86] Hans Hellmann. *Hans Hellmann: Einführung in Die Quantenchemie: Mit Biografischen Notizen von Hans Hellmann Jr.* Springer-Verlag, 2015.

[87] Richard Phillips Feynman. “Forces in molecules.” In: *Physical review* 56.4 (1939), p. 340.

[88] Stefano Baroni et al. “Phonons and related crystal properties from density-functional perturbation theory.” In: *Reviews of modern Physics* 73.2 (2001), p. 515.

[89] Pablo Ramos and Michele Pavanello. “Low-lying excited states by constrained DFT.” In: *The Journal of Chemical Physics* 148.14 (2018).

[90] Stefano de Gironcoli. *Lecture Notes on Relaxation Methods*. 2024. URL: https://people.sissa.it/~degironc/ES/lectures/Relax.pdf.

[91] Charles George Broyden. "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.

[92] Roger Fletcher. "A new approach to variable metric algorithms." In: *The computer journal* 13.3 (1970), pp. 317–322.

[93] Donald Goldfarb. "A family of variable-metric methods derived by variational means." In: *Mathematics of computation* 24.109 (1970), pp. 23–26.

[94] David F Shanno. "Conditioning of quasi-Newton methods for function minimization." In: *Mathematics of computation* 24.111 (1970), pp. 647–656.

[95] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials." In: *Journal of physics: Condensed matter* 21.39 (2009), p. 395502.

[96] Norman Troullier and José Luís Martins. "Efficient pseudopotentials for plane-wave calculations." In: *Physical review B* 43.3 (1991), p. 1993.

[97] David J Willock. *Molecular symmetry*. John Wiley & Sons, 2009.

[98] National Institute of Standards and Technology. *Carbon Monoxide - Ion Energetics*. 2023. URL: `https://webbook.nist.gov/cgi/cbook.cgi?ID=C630080&Mask=20#Ion-Energetics`.

[99] James R Chelikowsky, Leeor Kronik, and Igor Vasiliev. "Time-dependent density-functional calculations for the optical spectra of molecules, clusters, and nanocrystals." In: *Journal of Physics: Condensed Matter* 15.35 (2003), R1517.

[100] Onur Çaylak and Björn Baumeier. "Excited-state geometry optimization of small molecules with Many-Body Green's Functions Theory." In: *Journal of Chemical Theory and Computation* 17.2 (2021), pp. 879–888.

[101] JF Stanton. "Gauss, JR Ishikawa, N. Head-Gordon, M. 3." In: *Chem. Phys* 103.10 (1995), p. 4160.

[102] Gerhard Herzberg. *Molecular spectra and molecular structure*. D. van Nostrand, 1945.

[103] Barbara J Finlayson-Pitts and James N Pitts Jr. "Atmospheric chemistry. Fundamentals and experimental techniques." In: (1986).

[104] S Gomez-Carrasco, T Muller, and H Koppel. "Ab initio study of the VUV-induced multistate photodynamics of formaldehyde." In: *The Journal of Physical Chemistry A* 114.43 (2010), pp. 11436–11449.

[105] G Wilse Robinson and V Erdmanis DiGiorgio. "The nature of formaldehyde in its low-lying excited states." In: *Canadian Journal of Chemistry* 36.1 (1958), pp. 31–38.

[106] Jacob Kongsted et al. "The n→ $\pi^*$ electronic transition in microsolvated formaldehyde. A coupled cluster and combined coupled cluster/molecular mechanics study." In: *The Journal of Physical Chemistry A* 108.41 (2004), pp. 8624–8632.

[107] Monika Dash et al. "Excited states with selected configuration interaction-quantum Monte Carlo: Chemically accurate excitation energies and geometries." In: *Journal of chemical theory and computation* 15.9 (2019), pp. 4896–4906.

[108] Pauli Virtanen et al. "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python." In: *Nature Methods* 17 (2020), pp. 261–272. DOI: 10.1038/s41592-019-0686-2.

[109] Stephen J Wright. *Numerical optimization*. 2006.

[110] H Hellmann. *„Einführung in die Quantenchemie "Deuticke*. 1937.

[111] Francisco M Fernández. "Comment on "Breakdown of the Hellmann-Feynman theorem: Degeneracy is the key"." In: *Physical Review B* 69.3 (2004), p. 037101.