

Universidade de São Paulo
Instituto de Física

Estudo teórico de oligômeros de tiofeno para
fotovoltaicos: tiofeno-furano para filmes e tiofeno
cianoacrilado como sensibilizador de ZnO

Vinicius Alves Bastos

Orientadora: Prof.^a Dr.^a Marília Junqueira Caldas

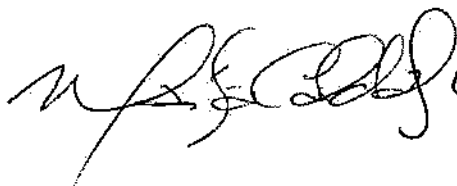
Dissertação de mestrado apresentada ao Instituto de Física
da Universidade de São Paulo, como requisito parcial para
a obtenção do título de Mestre em Ciências.

Banca Examinadora:

Prof.^a Dr.^a Marília Junqueira Caldas (IFUSP)

Prof. Dr. Rodrigo Barbosa Capaz (LNNano)

Prof. Dr. Ivan Helmuth Bechtold (UFSC)



São Paulo

2023

FICHA CATALOGRÁFICA
Preparada pelo Serviço de Biblioteca e Informação
do Instituto de Física da Universidade de São Paulo

Bastos, Vinicius Alves

Estudo teórico de oligômeros de tiofeno para fotovoltaicos: tiofeno-furano para filmes e tiofeno cianoacrilado como sensibilizador de ZnO. São Paulo, 2023.

Dissertação (Mestrado) - Universidade de São Paulo. Instituto de Física. Depto. de Física dos Materiais e Mecânica.

Orientador: Profa. Dra. Marília Junqueira Caldas.

Área de Concentração: Física.

Unitermos: 1. Física da matéria condensada; 2. Física do estado sólido; 3. Fotodetectores.

USP/IF/SBI-070/2023

University of São Paulo
Physics Institute

Theoretical study of thiophene oligomers for
photovoltaics: thiophene-furan for films and
cyanoacrylated thiophene as sensitizer for ZnO

Vinicius Alves Bastos

Supervisor: Prof. Dr. Marília Junqueira Caldas

Dissertation submitted to the Physics Institute of the University of São Paulo in partial fulfillment of the requirements for the degree of Master of Science.

Examining Committee:

Prof. Dr. Marília Junqueira Caldas (IFUSP)

Prof. Dr. Rodrigo Barbosa Capaz (LNNano)

Prof. Dr. Ivan Helmuth Bechtold (UFSC)

São Paulo

2023

Dedicatória

Aos meus pais Marta dos Santos Alves Bastos e Vicená Braga Bastos que apesar das circunstâncias, nunca mediram esforços para que eu e minha irmã tivéssemos uma boa educação, à minha irmã Danielli Alves Bastos que sempre me inspirou acerca das oportunidades que o estudo e dedicação me trariam e à minha companheira Chaiene Buchdid que esteve comigo nas horas mais difíceis e nos momentos mais alegres.

Agradecimentos

Meus mais sinceros agradecimentos à Professora Doutora Marília Junqueira Caldas que tem me orientado e auxiliado na minha formação científica desde o início da minha graduação em Física. Ao Doutor Tales José da Silva pelas valiosas conversas, por todo o conhecimento passado e pela colaboração em todas as etapas deste trabalho. À Sandra Regina Rodrigues Ribeiro secretária do DFMT por ser sempre tão prestativa e me ajudar com as questões burocráticas. À minha companheira Chaiene Buchdid por todo apoio e paciência durante todo o decorrer desta pesquisa. Aos meus pais Marta dos Santos Alves Bastos e Vicená Braga Bastos, minha irmã Danielli Alves Bastos e meu cunhado Fernando Gonçalves de Araujo por todo o apoio e pelas conversas de incentivo. Ao meu sogro José Eduardo Buchdid e minha sogra Luzia Pereira Buchdid por todo apoio e amizade. A todos os meus demais familiares e amigos pelo carinho e apoio sem os quais este trabalho não seria possível. Em especial eu agradeço: Robson, Vilma, Gerônimo, Henri e Rafaela Trevizol que apesar da distância neste período são para mim como minha família. Também agradeço à Gabriela T. Moreno de Sousa pela amizade e pelos momentos de desabafo.

O presente trabalho foi realizado com apoio financeiro do CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brasil, e do Instituto Nacional de Eletrônica Orgânica (INCT-INEO). Foram utilizados os parques computacionais do HPC-USP e CENAPAD-SP.

“The most beautiful thing that we can experience is the mysterious. It is the source of all true art and all science. He to whom this emotion is a stranger, who can no longer pause to wonder and stand rapt in awe, is as good as dead; his eyes are closed.”

— Albert Einstein

Abstract

Thiophene oligomers (oligothiophenes) are important organic materials for the photogeneration of electrical energy. Still, there are some issues concerning their application in films, such as the non-planarity of the linear chains, affecting molecular packing. An interesting suggestion reported in the literature is the co-polymerization of thiophene and furan since furan oligomers (oligofurans) are highly soluble and show a planar backbone. Thus, for films, it would be interesting to use this thiophene-furan composition. In another direction, an important point is the photoinduced charge transfer from the organic region of a photovoltaic device to the semiconductor contact. Another recent suggestion was adopting short cyanoacrylated oligothiophenes (*n*TCs) as sensitizers in Dye-Sensitized Solar Cells (DSSCs) based on *ZnO*.

Here, we perform a theoretical investigation of thiophene (T) and furan (F) oligomers/co-oligomers with 1 to 4 units, and *n*TCs containing 1 and 3 thiophene units. We adopt MP2-corrected HF, DFT (PBE), and hybrid HF/DFT (PBE0) methodologies to study the structural and ground-state electronic structure properties and the GW method and the Bethe-Salpeter equation (BSE) for quasiparticle and optical excitations. We show that avoiding TT sequencing in the thiophene-furan chains leads to planarity. Furthermore, thiophene-furan co-oligomers incorporate good electronic and optical properties of both thiophene and furan precursors: IPs similar to oligofurans and HOMO-LUMO and optical gaps similar to oligothiophenes. Regarding the *n*TCs, they present a donor-acceptor character with the HOMO electronic density mainly on the thiophene part and the LUMO on the cyanoacrylate termination. Consequently, the IPs are very similar to those for the thiophene-only oligomers, and the HOMO-LUMO gaps are smaller than for the oligothiophenes. The first optical excitation is mainly HOMO→LUMO, and for 3TC is around 2.5eV, in a good range for sunlight harvesting.

We further performed a PBE/PBE0 investigation of a single 3TC attached to the *ZnO* (10 $\bar{1}$ 0) surface (3TC-*ZnO*). We see that the 3TC-HOMO is in the *ZnO* gap and does not communicate with the surface states, while the 3TC-LUMO is in the conduction band (CB) region, and hybridizes with the crystal states (charge-transfer states). The 3TC-*ZnO* setup thus really provides properties suitable for application in DSSCs.

Keywords: Thiophene, Furan, dye sensitizing ZnO, DFT, BSE/GW

Resumo

Oligômeros de tiofeno (oligotiofenos) são materiais orgânicos importantes para a fotogeração de energia elétrica. Ainda assim, existem alguns problemas relativos à sua aplicação em filmes, como a não planaridade das cadeias lineares, afetando o empacotamento molecular. Uma proposta interessante sugerida na literatura é a copolimerização de tiofeno e furano, uma vez que os oligômeros de furano (oligofuranos) são altamente solúveis e apresentam estruturas planares. Assim, para filmes seria interessante utilizar essa composição tiofeno-furano. Em outra direção, um ponto importante é a transferência eletrônica fotoinduzida da parte orgânica para o contato semiconductor. Aqui, outra sugestão recente foi adotar oligotiofenos cianoacrilados curtos (*n*TCs) como sensitizadores em células solares sensibilizadas por corantes (DSSCs) baseadas em *ZnO*.

Neste trabalho, realizamos uma investigação teórica de oligômeros/co-oligômeros de tiofeno (T) e furano (F) com 1 a 4 anéis e *n*TCs contendo 1 e 3 anéis de tiofeno. Adotamos as metodologias MP2, DFT (PBE) e híbrida HF/DFT (PBE0) para estudar as propriedades estruturais e de estrutura eletrônica do estado fundamental, e o método GW e a equação de Bethe-Salpeter (BSE) para excitações de quasipartícula e ópticas. Mostramos que evitar o sequenciamento de TT nas cadeias de tiofeno-furano as leva à planaridade. Além disso, os co-oligômeros de tiofeno-furano incorporam as propriedades eletrônicas e ópticas dos precursores tiofeno e furano: IPs semelhantes aos oligofuranos e gaps HOMO-LUMO e ópticos semelhantes aos oligotiofenos. Já os *n*TCs apresentam caráter doador-aceitador sendo o HOMO principalmente na parte de tiofeno e o LUMO na terminação cianoacrilato. Consequentemente, os IPs são muito semelhantes aos dos oligotiofenos, e os gaps HOMO-LUMO são menores do que para os oligotiofenos. Sua primeira excitação óptica é principalmente HOMO→LUMO, e para o 3TC, em torno de 2,5eV, em uma boa faixa para captação de luz solar.

Além disso, realizamos uma investigação PBE/PBE0 de um 3TC ligado a superfície de *ZnO* (10 $\bar{1}$ 0) (3TC-*ZnO*). Vemos que o 3TC-HOMO está no gap do *ZnO* e não se comunica com os estados de superfície, enquanto o 3TC-LUMO está na região da banda de condução (CB) e hibridiza com os estados do cristal (estados de transferência de carga). Portanto, o 3TC-*ZnO* realmente exhibe propriedades adequadas para aplicação em DSSCs.

Palavras-chave: Tiofeno, Furano, corante sensitizando *ZnO*, DFT, BSE/GW

Contents

1	Introduction	8
2	Methodology	11
2.1	Many-body problem	11
2.2	Mean-field theories	12
2.2.1	Hartree-Fock	12
2.2.2	Density Functional Theory	14
2.3	Many-Body Perturbation Theory (MBPT)	16
2.3.1	Møller-Plesset perturbation theory	16
2.3.2	GW Method	17
2.3.3	Bethe-Salpeter Equation	21
2.4	Hubbard-U correction	22
2.5	Basis sets	23
2.5.1	Linear Combination of Atomic Orbitals (LCAO)	24
2.5.2	Planewave basis set (PW)	28
2.6	Computational packages	31
2.6.1	NWChem code	32
2.6.2	FHI-aims code	32
2.6.3	Quantum Espresso code (QE)	33
3	Thiophene-based oligomers	34
3.1	Thiophene linear oligomers	35
3.2	Thiophene-furan co-oligomers	40
3.3	Cyanoacrylated oligothiophenes	47
4	Thiophene as dye sensitizer for solar cells	56
4.1	ZnO crystal and (10 $\bar{1}$ 0) surface	56
4.2	ZnO (10 $\bar{1}$ 0) surface functionalized with <i>n</i> TcS	62
5	Conclusions	69

A	Appendix	72
A.1	Bond lengths and bond angles of 3T and 4T through different methods	72

Introduction

In the last years, the entire world community has been mobilized around the global energy crisis. The world energy demand has rapidly increased with increasing population and industry demand in developing countries. An urgent challenge has been posed: developing new and efficient technologies for energy generation. On the other hand, we are also facing an environmental crisis as a consequence of the greenhouse effect with consequences already seen around the globe. Accordingly, humanity can no longer rely on fossil-based fuels as the main source of energy production and must adopt new strategies with less environmental impact.

Among all energy sources available, the sun is probably the most promising since it generates an enormous amount of energy that can be collected at the Earth's surface and is a clean energy source, which means that, in principle, it does not produce carbon emissions during electricity generation. A way of exploiting sunlight energy is through a photovoltaic device. In these devices, the sunlight is absorbed by an active medium through the photoexcitation of electrons from occupied electronic states to unoccupied states. In Organic Solar Cells (OSCs) [1], this medium is formed by organic molecules, and it is sandwiched between two metal electrodes (cathode and anode) of different work functions to create a driving force for collecting the photogenerated charges. The vacancies of electrons created during photoexcitation, called "holes", also transport charge that must be guided through the molecular medium to the anode.

In this regard, thiophene linear oligomers or oligothiophenes are important organic molecules that possess essential characteristics for application in the active medium of OSCs [2, 3, 4], namely an electronic gap in the semiconductor region, light harvesting in the visible range of the solar spectrum, good charge mobility, good stability at ambient conditions, consolidated synthetic routes, and more. Nonetheless, these materials display low solubility in common organic solvents [5, 6], then the usual practice for their processing is alkylating the chains to improve their solubility. On the other hand, the alkyl side chains generally reduce the crystallinity of the derived films and, consequently, the molecules π -stacking, which harms the charge mobility [7].

An interesting proposition is the co-polymerization of thiophene (T) and furan (F) to address the problems concerning oligothiophenes solubility and the crystallinity of the films [5, 6, 8, 9]. Furan oligomers (oligofurans) are another class of relevant materials for organic electronic

applications. They have been widely explored in the last few years since they exhibit similar to thiophene electronic structure properties and higher solubility rates [10, 11]. Additionally, they can be manufactured from biomass and are expected to be biodegradable materials [11, 12]. However, oligofurans display low stability in the presence of oxygen and light which can compromise device lifetime [9]. Thus, thiophene and furan co-polymerization can provide an enhanced class of materials combining the qualities of both precursors.

Indeed, Gidron *et al.* [6] have investigated two thiophene-furan tetramers: TFFT and FTTF, and found higher luminescence, solubility, and backbone rigidity for the bifuran (FF) containing compound. In another work by Qiu *et al.* [8], they have grown furan and thiophene-furan polymers through catalyst-transfer polycondensation and obtained better uniformity for the thiophene-furan polymer-based film. Numerous works in the literature report important structural, electronic, and optical properties modifications in molecules, brought on by substituting thiophene or furan portions by analogs of thiophene-furan [5, 13, 14, 15]. Nonetheless, there is a lack of theoretical works concerning the electronic structure and optical properties of the thiophene-furan building blocks, since more attention has been given to the investigation of their structural properties [6, 16, 17].

In this work, we perform a systematic theoretical investigation of the structural, electronic, and optical properties of thiophene and furan oligomers/co-oligomers up to 4 rings to understand the modifications brought on by thiophene-furan co-polymerization. We employ MP2-corrected Hartree-Fock (HF), Density Functional Theory (DFT), and hybrid HF/DFT methodologies to study their structural and ground-state electronic structure properties, and the GW method and the Bethe-Salpeter equation (BSE) for quasiparticle and optical excitations. We show that thiophene-furan co-polymerization improves upon the structural and light-absorption properties of short oligothiophenes.

In another direction, cyanoacrylate-terminated oligothiophene dyes (n TCs) have a potential application as sensitizers in Dye-Sensitized Solar Cells (DSSCs) [18, 19, 20, 21], especially those adopting a molecular medium as the hole-transporting layer. In the active layer of these devices, the dyes are attached to metal-oxide nanocrystals (electron-transporting layer) through the terminal group, and these functionalized nanostructures are immersed in a molecular medium that acts as the hole-transporting layer. The n TCs are reported to present a donor-acceptor character since the cyanoacrylate termination acts as a strong electron-withdrawing group allowing photoexcited charge transfer from the thiophene part to the cyanoacrylate termination [18, 20, 21]. This mechanism, in turn, can favor photoexcited molecule-crystal charge injection in the attached systems, depending on the anchoring group and the molecule-surface energy levels alignment [18, 22, 23].

Abate *et al.* [18] have studied DSSCs adopting mesoporous TiO_2 films sensitized with cyanoacrylated pentathiophene molecules and reported good power conversion efficiency. In the following work [19], they have studied the application of cyanoacrylated oligothiophene molecules, with the number of thiophene units ranging from 1 to 5, as in the same previous setup

but adopting P3HT as the hole-transporting layer and light-absorber. In this case, they reported some problems concerning charge extraction, namely TC creates a barrier for charge collection from P3HT and 5TC cannot effectively inject charge into TiO_2 . Thus, they only recommend the use of 3TC and 4TC. More recently, Oehrlein *et al.* have reported fluorescence quenching of cyanoacrylated oligothiophene dyes (2TC to 5TC) upon binding to ZnO nanocrystals, indicating molecule-crystal charge transfer. They obtain similar time constants for the excited-state charge transfer of all studied molecules and ascribe to similar LUMO energy values.

Motivated by these findings we perform a theoretical investigation of the structural, electronic, and optical properties of cyanoacrylated oligothiophenes with 1 and 3 thiophene units, where for this last we studied two suggested structures which adopt different alkylation schemes [21, 24]. We employ DFT and hybrid HF/DFT methodologies to study the structural and ground-state electronic structure properties, and the GW method and the BSE for quasiparticle and optical excitations. We have further carried out a DFT and hybrid HF/DFT study of these molecules attached to the ZnO ($10\bar{1}0$) surface to investigate the molecule-crystal charge injection mechanism. This surface of the wurtzite ZnO crystal is particularly interesting since it is the most stable and the dominant surface in many ZnO nanostructures [25, 26]. Additionally, the ZnO crystal has been shown to be a promising material for the next generation of devices [27, 28]. We confirm that 3TC displays desirable properties for the application as a dye sensitizer in DSSCs and exhibits probable exciton dissociation at the 3TC- ZnO interface.

In the next chapters, we start with a short description of the theoretical methods adopted here (second chapter) and then move on to our results. In the third chapter, we present a discussion of the thiophene-based oligomers, starting with the investigation of the pure oligothiophenes and then the properties of thiophene-furan co-oligomers and cyanoacrylated oligothiophenes regarding their application as active medium in OSCs. In the fourth chapter, we present a study of a cyanoacrylated terthiophene molecule anchored to the ZnO ($10\bar{1}0$) surface, focusing on the photoexcited molecule-crystal charge transfer. Finally, in the fifth chapter, we present the conclusions and final remarks of this work.

Methodology

In this work, we have studied molecules and solid-state materials using consolidated theoretical methodologies in materials physics. The ground-state electronic structure properties were investigated using HF and/or DFT, and the excited-state properties were obtained through GW perturbation theory and the Bethe-Salpeter Equation. Some other methodologies such as the Møller-Plesset perturbation theory and the DFT+U, *i.e.*, DFT with a Hubbard potential correction, were also adopted and will be discussed in this chapter. Beyond that, the implementation of all these methods requires a suitable framework that is supported by different computational packages. These codes can be based on a Linear Combination of Atomic Orbitals (LCAO) or on Planewaves expansion (PW), among other basis set methodologies adopted for the representation of basic physical quantities, *e.g.*, the many-electron Hamiltonian, the wavefunctions, the electronic charge density, etc. Accordingly, the choice of methodology will depend on the kind of system addressed, each one having its own specificities. Here, we further present a brief discussion about the specific codes adopted in this work and their respective supported frameworks.

2.1 Many-body problem

For a system composed of N_n nucleus and N_e electrons, disregarding relativistic effects, the Hamiltonian (\hat{H}) can be expressed in Hartree atomic units (\hbar , e , m_e , and k_e are all unity) as

$$\hat{H} = -\frac{1}{2} \sum_{I=1}^{N_n} \frac{\nabla_{\vec{R}_I}^2}{M_I} - \frac{1}{2} \sum_{i=1}^{N_e} \nabla_{\vec{r}_i}^2 - \sum_{I=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_I}{\|\vec{R}_I - \vec{r}_i\|} + \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{\|\vec{r}_i - \vec{r}_j\|} + \sum_{I=1}^{N_n} \sum_{J>I}^{N_n} \frac{Z_I Z_J}{\|\vec{R}_I - \vec{R}_J\|}, \quad (2.1)$$

the uppercase labels I and J concern the nuclei, and the lowercase labels i and j , the electrons. From left to right, the terms account for the kinetic energy of nuclei, the kinetic energy of electrons, the electron-nucleus Coulomb interactions, the electron-electron Coulomb interactions, and the nucleus-nucleus Coulomb interactions.

In this work, we are interested in phenomena in which the effect of the nuclei dynamics on the electronic wavefunction can be neglected (the Born-Oppenheimer approximation). In this

case, the nucleus-nucleus interactions are in general treated classically, *i.e.*, the nuclei Coulomb potential is calculated for a fixed atomic configuration and added to the total energy of the system. The main part is to solve the electronic Hamiltonian, here called \hat{H}_e , which is written as

$$\hat{H}_e = -\frac{1}{2} \sum_{i=1}^{N_e} \nabla_{r_i}^2 - \sum_{I=1}^{N_n} \sum_{i=1}^{N_e} \frac{Z_I}{\|\vec{R}_I - \vec{r}_i\|} + \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{\|\vec{r}_i - \vec{r}_j\|}, \quad (2.2)$$

and as we see, it depends on the set of nuclei positions $\{\vec{R}_I\}$. In our case, the set of positions $\{\vec{R}_I\}$ will be determined through geometry optimization, as will be discussed later. The solution to \hat{H}_e gives us the set of eigenvalues $\{E_n\}$ and eigenfunctions $\{\Phi_n(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)\}$ (\vec{x}_i coordinates of position and spin) concerning the multielectron states of the system.

Nonetheless, \hat{H}_e cannot be solved exactly apart from atoms of one electron as the hydrogen atom and hydrogenic species. Then, it is necessary to resort to approximations that will be the subject addressed in the next sections.

2.2 Mean-field theories

In this section, we present a discussion about the HF and the DFT theories which provide a feasible solution for the ground-state of a many-electron system dictated by a Hamiltonian as presented in Eq. 2.2. These theories are the fundamental implementations in many computational packages for materials physics simulations and they can be applied to obtain structural and ground-state electronic structure properties for molecules and solid-state materials.

2.2.1 Hartree-Fock

In Hartree-Fock (HF), the mean-field approximation and the Pauli exclusion principle are incorporated to solve the ground-state (GS) electronic structure of a many-electron system. In the mean-field approximation, instead of considering the multielectron character of the Hamiltonian \hat{H}_e in Eq. 2.2, we assume that each electron will interact with the electronic field due to all the electrons in the system. In practice, also accounting to the Pauli exclusion principle or anti-symmetry of the electronic wavefunction, the GS all-electron wavefunction $\Phi_0(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ is written as

$$\Phi_0(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(\vec{x}_1) & \phi_1(\vec{x}_2) & \cdots & \phi_1(\vec{x}_N) \\ \phi_2(\vec{x}_1) & \phi_2(\vec{x}_2) & \cdots & \phi_2(\vec{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(\vec{x}_1) & \phi_N(\vec{x}_2) & \cdots & \phi_N(\vec{x}_N) \end{vmatrix}, \quad (2.3)$$

a Slater determinant of one-electron states [29, 30]. As we see, in HF, $\Phi_0(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ is written as a sum of products of one-electron states $\{\phi_i(\vec{x}_j)\}$, as in a system of non-interacting electrons. In fact, we are considering the interactions between the electrons through an effec-

tive potential raised by the full electronic charge density of the system and the effect of the antisymmetry of the all-electron wavefunction.

The HF energy E_{HF} will be a functional of the one-electron states $\{\phi_i\}$ and, following Ref. [31], it can be written as

$$E_{\text{HF}} = \langle \Phi_0 | \hat{H}_e | \Phi_0 \rangle = \sum_{i=0}^{N_e} h_{ii} + \frac{1}{2} \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \langle ij || ij \rangle. \quad (2.4)$$

h_{ii} , also known as one-electron integrals, are the diagonal elements $\int d\vec{x} \phi_i^*(\vec{x}) h(\vec{r}) \phi_i(\vec{x})$ of the one-electron operator $h(\vec{r})$ given as

$$h(\vec{r}) = -\frac{\nabla_r^2}{2} + v_{\text{ext}}(\vec{r}), \quad (2.5)$$

where $v_{\text{ext}}(\vec{r})$ is the external potential due to the nuclei. $\langle ij || ij \rangle$ are the antisymmetrized two-electron integrals given as

$$\langle ij || ij \rangle = \langle ij | ij \rangle - \langle ij | ji \rangle, \quad (2.6)$$

where $\langle ij | kl \rangle$ is the Dirac notation for the two-electron integrals defined as

$$\langle ij | kl \rangle = \iint d\vec{x}_1 d\vec{x}_2 \frac{\phi_i^*(\vec{x}_1) \phi_j^*(\vec{x}_2) \phi_k(\vec{x}_1) \phi_l(\vec{x}_2)}{\|\vec{r}_2 - \vec{r}_1\|}. \quad (2.7)$$

To determine the set of $\{\phi_i\}$ states that minimize E_{HF} , *i.e.*, the GS of the multielectron system, we apply the variational method constrained to the condition of orthonormalization of the single-particle states $\langle \phi_i | \phi_j \rangle = \delta_{ij}$. We must determine the stationary points of the Lagrange equation [32] of the system, in other words, to solve

$$\delta \left\{ E_{\text{HF}}[\{\phi_i\}] - \sum_{ij} \varepsilon_{ij} (\langle \phi_i | \phi_j \rangle - \delta_{ij}) \right\} = 0, \quad (2.8)$$

whose variation must be carried out with respect to the $\{\phi_i\}$ states (ε_{ij} are Lagrange multipliers). The solution to Eq. 2.8 leads us to the set of Hartree-Fock equations of the system, written as

$$f(\vec{x}) \phi_i(\vec{x}) = \varepsilon_i \phi_i(\vec{x}), \quad f(\vec{x}) = h(\vec{r}) + v_{\text{HF}}(\vec{x}) \quad (2.9)$$

where $f(\vec{x})$ is the one-electron HF Hamiltonian, and $v_{\text{HF}}(\vec{x})$ is the HF potential defined as

$$v_{\text{HF}}(\vec{x}) = \sum_{j=1}^N \left\{ J_j(\vec{x}) - K_j(\vec{x}) \right\}. \quad (2.10)$$

The first term in braces in Eq. 2.10 is the Coulomb or Hartree operator defined as

$$J_j(\vec{x}_1) = \int d\vec{x}_2 \frac{\phi_j^*(\vec{x}_2) \phi_j(\vec{x}_2)}{\|\vec{r}_2 - \vec{r}_1\|}, \quad (2.11)$$

which is the Coulomb potential due to the charge density created by an electron at the eigenstate j . The second term is the Exchange or Fock operator and its definition is given in terms of its action on the one-electron states as

$$K_j(\vec{x}_1) \phi_i(\vec{x}_1) = \int d\vec{x}_2 \frac{\phi_j^*(\vec{x}_2) \phi_i(\vec{x}_2)}{\|\vec{r}_2 - \vec{r}_1\|} \phi_j(\vec{x}_1). \quad (2.12)$$

This term arises from the antisymmetry of the all-electron wavefunction and, differently from the Hartree term, is a non-local potential in the sense that its action depends on ϕ_i across all space, not only at the position \vec{x}_1 .

It can be shown that the HF energy, defined in Eq. 2.4, can be rewritten in terms of the eigenvalues ε_i in Eq. 2.9 as

$$E_{\text{HF}} = \sum_{i=1}^{N_e} \varepsilon_i - \frac{1}{2} \sum_{i=1}^{N_e} \sum_{j=1}^{N_e} \langle ij || ij \rangle. \quad (2.13)$$

Now, the problem remains to obtain the set of eigenvalues and eigenfunctions associated with the one-electron HF Hamiltonian (Eq. 2.9). In practice, we expand the set of one-electron states $\{\phi_i(\vec{x})\}$ in Eq. 2.9, in a suitable basis set $\phi_i(\vec{x}) = \sum_k C_{ki} \tilde{\phi}_k(\vec{x})$. Then, left multiplying Eq. 2.9 by $\tilde{\phi}_l^*(\vec{x})$ and integrating over \vec{x} brings us to

$$\sum_k \left\{ \underbrace{\left(\int d\vec{x} \tilde{\phi}_l^*(\vec{x}) f(\vec{x}) \tilde{\phi}_k(\vec{x}) \right) C_{ki}}_{F_{lk}} \right\} = \varepsilon_i \sum_k \left\{ \underbrace{\left(\int d\vec{x} \tilde{\phi}_l^*(\vec{x}) \tilde{\phi}_k(\vec{x}) \right) C_{ki}}_{S_{lk}} \right\}, \quad (2.14)$$

a matrix equation written in terms of the coefficients C_{ki} . The terms in parentheses can be regarded as matrix elements: the term before the equal sign is the one-electron HF matrix and is generally represented as F_{lk} , and the term after the equal sign is the overlap matrix, normally represented as S_{lk} . These equations are known as Roothaan equations and are commonly presented in the matrix form $\mathbf{FC} = \mathbf{SC}\varepsilon$, where ε is a diagonal matrix of the coefficients ε_i . Its solution is achieved iteratively since the one-electron HF Hamiltonian \mathbf{F} depends on the \mathbf{C} coefficients through the one-electron wavefunctions. At this point, we obtain the expected information concerning the set of eigenvalues $\{\varepsilon_i\}$ and eigenfunctions $\{\phi_i(\vec{x})\}$.

2.2.2 Density Functional Theory

Density Functional Theory (DFT) [33] states that the energy of the GS of a many-electron system subjected to an external potential $v_{\text{ext}}(\vec{r})$ is completely determined by the electronic density $n(\vec{r})$ of the system. Accordingly, its great achievement is the reinterpretation of the GS energy E as a functional of $n(\vec{r})$, a three-variable function. This dependence is much simpler than in the wavefunction-based descriptions since the many-electron wavefunction depends on $3N_e$ variables (N_e is the number of electrons), thus a very complex problem to be solved even for small systems with dozens of electrons. Beyond that, Kohn and Sham [34] proposed a self-consistent solution for $E[n(\vec{r})]$, in the same spirit as the Hartree-Fock equations including, however, treatment for both exchange and correlation effects in terms of $n(\vec{r})$.

According to Kohn and Sham [34], $E[n]$ can be written as

$$E[n] = T_s[n] + \int d\vec{r} v_{\text{ext}}(\vec{r}) n(\vec{r}) + \frac{1}{2} \int d\vec{r} \int d\vec{r}' \frac{n(\vec{r}) n(\vec{r}')}{|\vec{r} - \vec{r}'|} + E_{\text{XC}}[n], \quad (2.15)$$

where $T_s[n]$ is the kinetic energy in a mean-field sense, *i.e.*, as in a system of non-interacting electrons, and all the many-electron effects are incorporated in the exchange-correlation functional $E_{XC}[n]$. From the variational principle, if $E[n]$ is the energy concerning the GS electronic density $n(\vec{r})$ of a certain system, the evaluation of E at any other electronic distribution $n'(\vec{r})$ would result in a higher energy value. Then, the GS electronic density $n(\vec{r})$ will be a stationary point of $E[n]$, constrained by the total number of electrons of the system N_e . With this in mind, we must determine the stationary points of the Lagrange equation of the system, *i.e.*, to solve the equation

$$\delta \left\{ E[n(\vec{r})] - \varepsilon \left[\int d\vec{r} n(\vec{r}) - N_e \right] \right\} = 0, \quad (2.16)$$

with respect to $n(\vec{r})$. This leads us to the minimum condition, given as

$$\int d\vec{r} \delta n(\vec{r}) \left\{ \frac{\delta T_s[n]}{\delta n(\vec{r})} + v_{\text{ext}}(\vec{r}) + \int d\vec{r}' \frac{n(\vec{r}')}{\|\vec{r} - \vec{r}'\|} + \frac{\delta E_{XC}[n]}{\delta n(\vec{r})} - \varepsilon \right\} = 0. \quad (2.17)$$

Now, to arrive at the Kohn-Sham equations, the density $n(\vec{r})$ is written in terms of an orthogonal set of N_e single-electron functions $\{\phi_i(\vec{r})\}$ [35] as

$$n(\vec{r}) = \sum_{i=1}^{N_e} \phi_i^*(\vec{r}) \phi_i(\vec{r}), \quad (2.18)$$

and its variation is evaluated with respect to these functions. The Kohn-Sham equations are finally given as

$$\left\{ \nabla^2 + v_{\text{ext}}(\vec{r}) + \int d\vec{r}' \frac{n(\vec{r}')}{\|\vec{r} - \vec{r}'\|} + v_{XC}(\vec{r}) \right\} \phi_i(\vec{r}) = \varepsilon_i \phi_i(\vec{r}), \quad (2.19)$$

where $v_{XC}(\vec{r})$ is the effective potential due to $E_{XC}[n]$ defined as

$$v_{XC}(\vec{r}) \equiv \frac{\delta E_{XC}[n(\vec{r})]}{\delta n(\vec{r})}. \quad (2.20)$$

As in the Hartree-Fock method, the Kohn-Sham equations must be expanded in a suitable basis set, and the solution will be achieved self-consistently.

The problem now remains to find out the functional form of $E_{XC}[n]$. In fact, there is no exact formulation for $E_{XC}[n]$, so what is done in practice is adopting an approximation with the correct behavior in certain important regimes. Kohn and Sham [34] proposed the form

$$E_{XC}^{LDA}[n] = \int d\vec{r} n(\vec{r}) \varepsilon_{XC}(n(\vec{r})), \quad (2.21)$$

where $\varepsilon_{XC}(n(\vec{r}))$ is the exchange-correlation energy per electron of a uniform electron gas of density n . This approximation, called Local-Density Approximation (LDA), will be valid in the regime of slowly varying electronic density $n(\vec{r})$, as in this case, n can be assumed as locally constant.

A further step in the definition of $E_{\text{XC}}[n]$ is considering local variations in the density $n(\vec{r})$ by also incorporating a dependence with $\nabla n(\vec{r})$. Exchange-correlation functionals in this category of approximation are called Generalized Gradient Approximation (GGA). A remarkable proposition is the PBE functional given as

$$E_{\text{XC}}^{\text{PBE}}[n] = \int d\vec{r} f(n(\vec{r}), \nabla n(\vec{r})) \quad (2.22)$$

in the non-spinpolarized case, where f is built to satisfy well-established conditions concerning the electronic density [36]. It has been successfully applied in the study of several systems from molecules to solid-state materials. Usually, the PBE functional provides fine results for the structural properties of the materials; however, the electronic structure properties, such as the electronic gap and ionization energies, which are more dependent on the many-electron effects, are not properly described. An extension of this method, the PBE0 functional [37], mixes 25% of Hartree-Fock exchange (Eq. 2.12) with PBE exchange (hybrid HF/DFT method) to have a better description of the exchange potential, which is, in general, the foremost dominant term of E_{XC} . The PBE0 is then defined as

$$E_{\text{XC}}^{\text{PBE0}} = E_{\text{XC}}^{\text{PBE}} + \frac{1}{4} \left(E_{\text{X}}^{\text{HF}} - E_{\text{X}}^{\text{PBE}} \right). \quad (2.23)$$

Although much more demanding computationally than the PBE, it significantly improves on the description of the electronic structure properties of molecules and semiconductor materials.

2.3 Many-Body Perturbation Theory (MBPT)

A standard practice in physics is starting from a simple model to describe a specific system, and then, refining this model if the required phenomenon cannot be captured. Following this prescription, starting from a mean-field ground-state electronic structure, we can resort to perturbation theory in order to capture higher-order correlation effects or to obtain excited-state properties such as ionization energies and optical excitations. In this section, we briefly discuss the Møller-Plesset perturbation theory, the GW method, and the Bethe-Salpeter Equation, which are perturbative approaches to studying such phenomena.

2.3.1 Møller-Plesset perturbation theory

The Møller-Plesset perturbation theory [31, 38, 39] is a particular case of time-independent perturbation theory, where the HF Hamiltonian, defined as

$$F(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{N_e}) = \sum_{i=1}^{N_e} f(\vec{x}_i), \quad (2.24)$$

that is the sum of the one-electron HF Hamiltonians presented in Eq. 2.9, is perturbed by a potential \hat{V} holding all correlation effects due to the all-electron Coulomb interaction. Accord-

ingly, \hat{V} is given by

$$V(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_{N_e}) = \sum_{i=1}^{N_e} \sum_{j>i}^{N_e} \frac{1}{|\vec{r}_i - \vec{r}_j|} - \sum_{i=1}^{N_e} v_{\text{HF}}(\vec{x}_i), \quad (2.25)$$

which is the all-electron Coulomb interaction minus the HF potential (Eq. 2.10) due to all the electrons.

From the discussion in Section 2.2.1, we notice that the GS all-electron wavefunction $|\Phi_0\rangle$ of the HF Hamiltonian (Eq. 2.24) will be, without approximations, the Slater determinant (Eq. 2.3) of the N_e lowest-energy one-electron wavefunctions $\{\phi_i\}$ (Eq. 2.9). Thus, the GS energy E_0 will be written as $E_0 = \sum_{i=1}^{N_e} \varepsilon_i$, and further including the first-order correction term due to \hat{V} , that is $\langle \Phi_0 | \hat{V} | \Phi_0 \rangle$, will lead to the HF energy E_{HF} defined in Eq. 2.13.

The second-order correction term, given as

$$E_0^{(2)} = - \sum_{n \neq 0} \frac{|\langle \Phi_0 | \hat{V} | \Phi_n \rangle|^2}{E_n^{(0)} - E_0^{(0)}}, \quad (2.26)$$

ouples the GS wavefunction $|\Phi_0\rangle$ with the excited-state wavefunctions $|\Phi_n\rangle$. In this last, occupied one-electron states of $|\Phi_0\rangle$, labeled here by i, j, k, \dots , are replaced by unoccupied states a, b, c, \dots ($|\Phi_{ijk\dots}^{abc\dots}\rangle$). It can be shown that the matrix elements $\langle \Phi_0 | \hat{V} | \Phi_n \rangle$ are non-null only between $|\Phi_0\rangle$ and double-excited states $|\Phi_{ij}^{ab}\rangle$. Additionally, the zeroth-order energy term $E_n^{(0)}$ of any doubled-excited state $|\Phi_{ij}^{ab}\rangle$ can be expressed in terms of E_0 as $E_0 + \varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j$, where ε_i and ε_j are the energies concerning the respective occupied states and ε_a and ε_b of the unoccupied states. We finally arrive at the MP2 correction to E_{HF} given as

$$E_c^{\text{MP2}} = -\frac{1}{4} \sum_{i,j}^{\text{occ}} \sum_{a,b}^{\text{vir}} \frac{|\langle ij || ab \rangle|^2}{\varepsilon_a + \varepsilon_b - \varepsilon_i - \varepsilon_j}, \quad (2.27)$$

where $\langle ij || ab \rangle$ are antisymmetrized two-electron integrals defined in Eq. 2.6.

One could proceed by including higher-order correction terms, *i.e.*, MP3, MP4, and so on. Such higher-order corrections are beyond the scope of this work and, thus, will not be discussed here.

2.3.2 GW Method

The ionization energies of a many-electron system, associated with photoelectron spectroscopy measurements, can be formally determined through the GW theory. Following Ref. [40, 41], the single-particle Green's function of an N-electron system prepared at the ground-state $|\Psi_0^N\rangle$ is defined as

$$G(\vec{x}t, \vec{x}'t') = -i \left\langle \Psi_0^N \left| \hat{T} \left[\hat{\psi}(\vec{x}t) \hat{\psi}^\dagger(\vec{x}'t') \right] \right| \Psi_0^N \right\rangle, \quad (2.28)$$

where \hat{T} is the time-ordering operator and $\hat{\psi}^\dagger(\vec{x}'t')$ and $\hat{\psi}(\vec{x}t)$ are the creation and annihilation field operators of second-quantization in the Heisenberg picture [30]. The single-particle

Green's function of Eq. 2.28 is defined so that $iG(\vec{x}t, \vec{x}'t')$ gives the probability amplitude of a particle (electron or hole) inserted at (\vec{x}', t') to propagate to (\vec{x}, t) . By including the closure relation

$$\sum_i |\Psi_i^{N\pm 1}\rangle \langle \Psi_i^{N\pm 1}| = \mathbb{I} \quad (2.29)$$

between the field operators in Eq. 2.28 and Fourier transforming from time to frequency domain, we arrive at the Lehmann representation of the single-particle Green's function

$$G(\vec{x}, \vec{x}'; \omega) = \sum_i \frac{\psi_i^{N+1}(\vec{x}) \psi_i^{N+1*}(\vec{x}')}{\omega - \varepsilon_i^{N+1} + i\eta} + \sum_i \frac{\psi_i^{N-1}(\vec{x}) \psi_i^{N-1*}(\vec{x}')}{\omega - \varepsilon_i^{N-1} - i\eta}. \quad (2.30)$$

The $\{|\Psi_i^{N\pm 1}\rangle\}$ in Eq. 2.29 must be regarded as all many-body eigenfunctions with $N \pm 1$ electrons, and therefore, they form a complete basis set at the subspace of the many-body wavefunctions with $N \pm 1$ electrons. $\psi_i^{N-1}(\vec{x})$ and $\psi_i^{N+1}(\vec{x})$ are many-body quantities regarded as quasiparticle eigenfunctions and are defined by

$$\psi_i^{N-1}(\vec{x}) = \langle \Psi_i^{N-1} | \hat{\psi}(\vec{x}) | \Psi_0^N \rangle \quad \psi_i^{N+1}(\vec{x}) = \langle \Psi_0^N | \hat{\psi}(\vec{x}) | \Psi_i^{N+1} \rangle, \quad (2.31)$$

and $\{\varepsilon_i^{N\pm 1}\}$ are the many-body excitation energies given by

$$\varepsilon_i^{N-1} = E_0^N - E_i^{N-1} \quad \varepsilon_i^{N+1} = E_i^{N+1} - E_0^N. \quad (2.32)$$

We see that the single-particle Green's function in Eq. 2.30 has poles at the many-body excitation energies $\varepsilon_i^{N\pm 1}$, which are the quantities accessed in direct and inverse photoelectron spectroscopy. While $\{\psi_i^{N\pm 1}(\vec{x})\}$ are indeed many-body quantities, in the case of a noninteracting electron system, they will be simply the one-electron eigenstates and $\varepsilon_i^{N\pm 1}$ their respective one-electron eigenvalues.

The single-particle Green's function must obey the equation of motion [42] given as

$$\left[\omega - h_0(\vec{x}) \right] G(\vec{x}, \vec{x}'; \omega) - \int d\vec{x}'' \Sigma(\vec{x}, \vec{x}''; \omega) G(\vec{x}'', \vec{x}'; \omega) = \delta(\vec{x}, \vec{x}'), \quad (2.33)$$

where the one-electron operator defined in Eq. 2.5 and the Hartree potential (Eq. 2.11) were rewritten as $h_0(\vec{x})$, given by

$$h_0(\vec{x}) = h(\vec{r}) + v_H(\vec{x}). \quad (2.34)$$

Σ , in the second term of the left-hand side of Eq. 2.33, is the self-energy operator which accounts for all many-body effects as defined below. Substituting the single-particle Green's function defined in Eq. 2.30 into the equation of motion 2.33, we arrive at the quasiparticle equation given as

$$h_0(\vec{x}) \psi_i(\vec{x}) + \int d\vec{x}' \Sigma(\vec{x}, \vec{x}'; \varepsilon_i) \psi_i(\vec{x}') = \varepsilon_i \psi_i(\vec{x}). \quad (2.35)$$

Here, the superscripts on the quasiparticle eigenfunctions and the associated excitation energies were dropped for simplification. The solution of the quasiparticle equation (Eq. 2.35)

provides the quasiparticle eigenfunctions $\{\psi_i(\vec{x})\}$ and their respective excitation energies $\{\varepsilon_i\}$ concerning a many-electron system.

The self-energy can be formally obtained by Hedin's set of integrodifferential equations [42, 40], whose self-consistent solution completely solves the many-electron problem. In practice, this set of equations cannot be solved, but they can be computationally iterated to some order to obtain good approximations. The most usual, named one-shot GW or G0W0 approximation, consists of the first iteration of Hedin's equations. In this case, they reduce to

$$\Gamma(12;3) = \delta(12)\delta(13) \quad (2.36)$$

$$\chi_0(12) = -iG_0(12)G_0(21) \quad (2.37)$$

$$W(12) = v(12) + \iint d3 d4 v(13)\chi_0(34)W(42) \quad (2.38)$$

$$\Sigma(12) = iG_0(12)W(1^+2) \quad (2.39)$$

$$G(12) = G_0(12) + \iint d3 d4 G_0(13)\Sigma(34)G(42), \quad (2.40)$$

where Γ is the vertex function, χ_0 is the irreducible polarizability, v is the two-electron Coulomb potential, and W the screened Coulomb potential. The numbers denote coordinates of position, spin, and time, *e.g.*, $1 = (\vec{r}_1, \sigma_1, t_1)$, whereas the plus-indexed coordinate $1^+ = (\vec{r}_1, \sigma_1, t_1 + \xi)$, $\xi \rightarrow 0^+$, is adopted to ensure the correct time ordering of the field operators at the single-particle Green's function. W , in Eq. 2.38, can be rewritten in terms of the inverse dielectric function ε^{-1} as

$$W(12) = \int d3 \varepsilon^{-1}(13)v(32), \quad (2.41)$$

where the dielectric function ε is given by

$$\varepsilon(12) = \delta(12) - \int d3 v(13)\chi_0(32). \quad (2.42)$$

Interestingly, in Eq. 2.41 it is shown that the many-body effect can be accounted as a dielectric medium, screening the quasiparticle interactions.

In practice, the G0W0 approximation is performed on top of a mean-field converged electronic structure, *i.e.*, calculated with DFT, by considering the self-energy Σ as a perturbation to the XC potential of the form

$$\Sigma(\vec{x}, \vec{x}'; \varepsilon_i) = v_{\text{XC}}(\vec{r})\delta(\vec{x} - \vec{x}'), \quad (2.43)$$

whose effect is accounted by a first-order correction in time-independent perturbation theory. This is, in many cases, a good approximation since DFT provides electronic structure properties in good agreement with experimental data. The quasiparticle (qp) excitation energies will be given by

$$\varepsilon_i^{\text{qp}} \approx \varepsilon_i^{\text{KS}} + \langle \phi_i^{\text{KS}} | \Sigma(\varepsilon_i^{\text{qp}}) - v_{\text{XC}} | \phi_i^{\text{KS}} \rangle, \quad (2.44)$$

whose solution can be achieved iteratively starting with $\varepsilon_i^{\text{qp}} = \varepsilon_i^{\text{KS}}$.

To evaluate the self-energy in Eq. 2.43, the definition given in Eq. 2.39 is Fourier transformed from time to frequency domain [41] leading to

$$\Sigma(\vec{x}, \vec{x}', \omega) = \frac{i}{2\pi} \int d\omega' e^{i\omega'\eta} G_0(\vec{x}, \vec{x}', \omega + \omega') W_0(\vec{x}, \vec{x}', \omega'), \quad (2.45)$$

where η is a positive infinitesimal number. G_0 is the single-particle Green's function for the DFT converged electronic structure [41] and is given by

$$G_0(\vec{x}, \vec{x}', \omega) = \sum_i \frac{\phi_i(\vec{x}) \phi_i^*(\vec{x}')}{\omega - \varepsilon_i - i\eta \operatorname{sgn}(E_F - \varepsilon_i)} \quad (2.46)$$

in terms of the Kohn-Sham set of eigenfunctions $\{\phi_i(\vec{x})\}$ and eigenvalues $\{\varepsilon_i\}$. Σ can be divided into a correlation part given by

$$\Sigma^C(\vec{x}, \vec{x}', \omega) = \frac{i}{2\pi} \int d\omega' G_0(\vec{x}, \vec{x}', \omega + \omega') W_0^C(\vec{x}, \vec{x}', \omega'), \quad (2.47)$$

where W_0^C is defined as

$$W_0^C(\vec{x}, \vec{x}', \omega) = W_0(\vec{x}, \vec{x}', \omega) - v(\vec{r}, \vec{r}'), \quad (2.48)$$

and an exchange part

$$\begin{aligned} \Sigma^X(\vec{x}, \vec{x}') &= \frac{i}{2\pi} \int d\omega' e^{i\omega'\eta} G_0(\vec{x}, \vec{x}', \omega + \omega') v(\vec{r}, \vec{r}') \\ &= \sum_i^{\text{occ}} \phi_i(\vec{x}) \phi_i^*(\vec{x}') v(\vec{r}, \vec{r}') \end{aligned} \quad (2.49)$$

The exchange part Σ^X is exactly the Hartree-Fock potential (Eq. 2.12), whereas the correlation part Σ^C can be recognized as a non-local correlation potential. The first is calculated as usual, whereas for the second, there are some numerical integration techniques described in the literature [41]. In the calculations performed using the FHI-aims code, we adopted the Analytic Continuation (AC) technique with a two-pole model expansion of the self-energy. Basically, the self-energy is integrated along the imaginary frequency axis given as

$$\Sigma^C(\vec{x}, \vec{x}', i\omega) = \frac{i}{2\pi} \int d\omega' G_0(\vec{x}, \vec{x}', i\omega + i\omega') W_0^C(\vec{x}, \vec{x}', i\omega'), \quad (2.50)$$

which is easier to evaluate than in the real frequency axis. To return from the imaginary to the real frequency axis, the self-energy is calculated at a set of imaginary frequencies $\{i\omega\}$ and interpolated by a two-pole model function through a nonlinear least-squares fitting given by

$$\Sigma_i^C(i\omega) \approx \sum_{j=1}^2 \frac{a_{i,j}}{i\omega + b_{i,j}} + a_{i,0}. \quad (2.51)$$

This interpolation is performed for each term $\langle \phi_i^{\text{KS}} | \Sigma(\varepsilon_i) | \phi_i^{\text{KS}} \rangle$ in Eq. 2.44, around $\varepsilon_i = \varepsilon_i^{\text{KS}}$, to determine the set of coefficients of the expansion. From the identity theorem of complex analysis, the model function can be extended from the imaginary to the real frequency axis. A more detailed discussion of this numerical method and others is given in Ref. [41] and in the manual of the FHI-aims code.

2.3.3 Bethe-Salpeter Equation

To obtain the optical-excitation energies, *e.g.*, accessed in absorption spectroscopy measurements, we adopted the Bethe-Salpeter equation (BSE) [41, 43, 44, 45, 46]. Following Ref. [41, 45, 46], it can be expressed in a matrix form written as

$$\begin{bmatrix} R & C \\ -C^* & -R^* \end{bmatrix} \begin{bmatrix} X^m \\ Y^m \end{bmatrix} = \Omega^m \begin{bmatrix} X^m \\ Y^m \end{bmatrix}, \quad (2.52)$$

in the product space basis of occupied (i, j) and unoccupied states (a, b) given by $\{\phi_i(\vec{r})\phi_a(\vec{r}')\}$. The matrix elements R and C are defined as

$$R_{ia,jb} = (\varepsilon_a^{\text{qp}} - \varepsilon_i^{\text{qp}})\delta_{ij}\delta_{ab} + \kappa(ia|jb) - W_{ij,ab}, \quad (2.53)$$

$$C_{ia,jb} = \kappa(ia|bj) - W_{ib,aj}, \quad (2.54)$$

where $\{\varepsilon_{i/a}^{\text{qp}}\}$ are the quasiparticle energies obtained in the GW calculations, κ is equal to 2 or 0 to address singlet or triplet excited states, respectively, and the so-called Mulliken notation [45, 46] is adopted in the definition of $(ia|jb)$ and $W_{ij,ab}$ given as

$$(ia|jb) = \iint d\vec{x}d\vec{x}' \frac{\phi_i^*(\vec{x})\phi_a(\vec{x})\phi_j^*(\vec{x}')\phi_b(\vec{x}')}{\|\vec{r} - \vec{r}'\|} \quad (2.55)$$

$$W_{ij,ab} = \iint d\vec{x}d\vec{x}' \phi_i^*(\vec{x})\phi_j(\vec{x})W(\vec{x},\vec{x}';\omega=0)\phi_a^*(\vec{x}')\phi_b(\vec{x}'). \quad (2.56)$$

$W(\omega=0)$ is the screened coulomb potential defined in the GW theory (Eq. 2.41) and evaluated in the static limit. $\{X^m, Y^m\}$ in Eq. 2.52 are the expansion coefficients of the electron-hole wavefunctions (eh) given as

$$\Psi_m^{\text{eh}}(\vec{r}_e, \vec{r}_p) = \sum_{ia} [X_{ia}^m \phi_i(\vec{r}_h)\phi_a(\vec{r}_e) + Y_{ia}^m \phi_i(\vec{r}_e)\phi_a(\vec{r}_h)], \quad (2.57)$$

and $\{\Omega^m\}$ their respective excitation energies. In the FHI-aims code, the BSE implementation adopts the Tamm-Dancoff Approximation (TDA) that consists of neglecting the coupling term (C) between the resonant (R) and antiresonant ($-R^*$) parts of the matrix form in Eq. 2.52.

The oscillator strength f_m concerning the optical excitation m can be calculated with the respective coefficients $\{X^m, Y^m\}$ and excitation energy Ω^m obtained as solution for the BSE eigenvalue problem [46] as

$$f_m = \frac{2}{3}\Omega^m \sum_{\mu=x,y,z} (d_{m,\mu})^2, \quad (2.58)$$

with $d_{m,\mu}$, the transition dipole moment along the direction μ ($\mu = x, y, z$), given as

$$d_{m,\mu} = \sqrt{2} \sum_{i,a} \{ \langle \phi_i | \hat{\mu} | \phi_a \rangle (X_{ia}^m + Y_{ia}^m) \}. \quad (2.59)$$

2.4 Hubbard-U correction

DFT usually does not correctly describe d and f states [47, 48], then a correction is advised known as Hubbard potential correction. For example, the Zn $3d$ -states in the wurtzite ZnO crystal have higher energy values in DFT calculations than the experimental reference values [49] with respect to the VBT. Consequently, in DFT calculations these states are hybridized with the O $2p$ -states around the VBT in disagreement with the experimental results. In the work by Li *et al.* [50], they show that by adopting a Hubbard-U treatment for the Zn $3d$ -states ($U=8.5\text{eV}$) places the $3d$ -band at an energy value of -7.7eV with respect to the VBT in good agreement with the experimental value of -7.5eV .

Following Ref. [48], the effect of the Hubbard interaction can be illustrated through its formulation in a one-band model given as

$$\hat{H}_{\text{Hub}} = t \sum_{\langle i,j \rangle, \sigma} (\hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} + \hat{c}_{j,\sigma}^\dagger \hat{c}_{i,\sigma}) + U \sum_i n_{i,\uparrow} n_{i,\downarrow}, \quad (2.60)$$

in terms of creation and annihilation operators of second quantization [30]. $\langle i, j \rangle$ stand for nearest-neighbor atomic sites, σ is the spin angular momentum projection, and $n_{i,\uparrow}$ is the electronic density at the atomic site i concerning the spin channel \uparrow . The first term, called as ‘‘hopping’’ term, can be regarded as a tight-binding interaction, where t is the overlap integral between nearest-neighbor atomic sites that will give rise to the bandwidth. The second term is a hard shift that depends on the product of the electronic densities per spin channel at each atomic site. It accounts for Coulombic repulsive interactions between electrons at the same site. When $t \gg U$, the electron will behave as in a crystalline structure, whereas in the regime of $U \gg t$, it will behave as an isolated atomic state even at half-filled bands. The first regime is well described in DFT, whereas for the second, we must rely on a correction for the Hubbard interaction.

The so-called LDA+U model works as regular DFT adopting the Hubbard Hamiltonian to treat the interactions between localized d or f electrons. In this case, the total energy functional is written as [48]

$$E_{\text{LDA+U}}[n(\vec{r})] = E_{\text{LDA}}[n(\vec{r})] + E_{\text{Hub}}[\{n_{mm'}^{I\sigma}\}] - E_{\text{dc}}[\{n^{I\sigma}\}], \quad (2.61)$$

where $E_{\text{Hub}}[\{n_{mm'}^{I\sigma}\}]$ is the energy concerning the Hubbard Hamiltonian and $E_{\text{dc}}[\{n^{I\sigma}\}]$ is a double-counting term to counteract the Hubbard energy accounted in DFT. $n_{mm'}^{I\sigma}$ is defined as

$$n_{mm'}^{I\sigma} = \sum_{n,\vec{k}} f(\epsilon_{n\vec{k}}^\sigma) \langle \psi_{n\vec{k}}^\sigma | \phi_m^I \rangle \langle \phi_m^I | \psi_{n\vec{k}}^\sigma \rangle, \quad (2.62)$$

where $\{|\phi_m^I\rangle\}$ are localized states at the atomic site I (d -type or f -type atomic orbitals) and $f(\epsilon_{n\vec{k}}^\sigma)$ is the electronic occupation of the state $|\psi_{n\vec{k}}^\sigma\rangle$ (Eq. 2.87), and $n^{I\sigma}$ is given by

$$n^{I\sigma} = \sum_m n_{mm}^{I\sigma}. \quad (2.63)$$

$n_{mm'}^{I\sigma}$ can be recognized as matrix elements of the density operator of the system in the subspace of the localized states at the atomic site I . Therefore, the sum elements in Eq. 2.63 ($\{n_{mm'}^{I\sigma}\}$) are the electronic populations of the $\{|\phi_m^I\rangle\}$ states. In its simplest formulation, the LDA+U energy functional is written as

$$E_{\text{LDA+U}} = E_{\text{LDA}} + \sum_I \left[\frac{U^I}{2} \left(\sum_{\{m,\sigma\} \neq \{m',\sigma'\}} n_m^{I\sigma} n_{m'}^{I\sigma'} \right) - \frac{U^I}{2} n^I (n^I - 1) \right], \quad (2.64)$$

where $n_m^{I\sigma} \equiv n_{mm}^{I\sigma}$ and $n^I = n^{I\uparrow} + n^{I\downarrow}$ (Eq. 2.63). The second term of the right-hand side of this equation is the sum of the Hubbard and the double-counting energies. Using the definition given in Eq. 2.62, it can be shown that the action of the LDA+U potential on the electronic states $\{|\psi_{nk}^\sigma\rangle\}$ will be given as

$$V_{\text{LDA+U}} |\psi_{nk}^\sigma\rangle = V_{\text{LDA}} |\psi_{nk}^\sigma\rangle + \sum_{I,m} \left[U^I \left(\frac{1}{2} - n_m^{I\sigma} \right) |\phi_m^I\rangle \langle \phi_m^I| \right] |\psi_{nk}^\sigma\rangle. \quad (2.65)$$

As we see, the Hubbard correction is positive for $|\psi_{nk}^\sigma\rangle$ states with less than half-filled atomic orbitals $\{|\phi_m^I\rangle\}$, whereas it is negative otherwise. This is the mechanism through which the Hubbard correction penalizes semi-occupied states which can hybridize with near-energy states.

In fact, we used the simplification of the rotationally invariant formulation proposed in the work by Liechtenstein *et al.* [51], since the above-presented formulation has an undesirable dependence on the choice of the localized basis set adopted for the d -type orbitals. In this case, the Hubbard functional is defined as [48]

$$\begin{aligned} E_{\text{U}}[\{n_{mm'}^{I\sigma}\}] &= E_{\text{Hub}}[\{n_{mm'}^{I\sigma}\}] - E_{\text{dc}}[\{n^{I\sigma}\}] \\ &= \sum_{I,\sigma} \frac{U^I}{2} \text{Tr}\{\mathbf{n}^{I\sigma}(1 - \mathbf{n}^{I\sigma})\}, \end{aligned} \quad (2.66)$$

where the symbol in bold is the occupation matrix defined in Eq. 2.62. This correction is implemented in the QE code.

2.5 Basis sets

As commented above, the presented physical theories need a suitable framework to be applied. In particular, a representation framework is required, *i.e.*, a set of basis functions upon which the physical quantities (Hamiltonian operator, wavefunctions, electronic density, etc.) will be represented. Among the computational methods available in the literature, LCAO-based methods are more convenient for the treatment of isolated systems, such as molecules and clusters, whereas PW-based methods are more convenient for treating periodic systems, *i.e.*, materials with a crystalline structure. It is worth noting that both basis sets can be applied in the two cases. In this section, we present a discussion about the basis set methodologies adopted in this work.

2.5.1 Linear Combination of Atomic Orbitals (LCAO)

As discussed in Section 2.2.1, we need a set of basis functions to represent the one-electron states $\{\phi_i(\vec{x})\}$. For a matter of convenience, this basis set will depend on the type of system addressed, *e.g.*, in the case of an isolated molecule an intuitive basis set would be the set of atomic orbitals of the atomic species present in the molecule. In fact, this reasoning is adopted in the so-called Linear Combination of Atomic Orbitals (LCAO) method.

Nonetheless, even isolated atoms except for hydrogen are many-electron systems whose atomic orbitals will be obtained in a basis set representation. However, disregarding the many-body effects and taking into account the spherical symmetry of the atoms, we can approximate the electrons as confined in a spherically symmetric effective potential defined as $v_{eff}(r)$. The eigenstates for such a class of systems [52, 53] will be given as

$$\varphi_{n,l,m}(r, \theta, \phi) = \frac{1}{r} u_{n,l}(r) Y_l^m(\theta, \phi), \quad (2.67)$$

where $\{u_{n,l}(r)\}$ are the solutions to the radial part of the electronic Hamiltonian and $\{Y_l^m(\theta, \phi)\}$, *i.e.*, the spherical harmonics, are the solutions to the angular part of the Laplace operator. l is an integer number higher or equal to zero, which determines the magnitude of the angular momentum operator \hat{L} , and $m = -l, -l+1, \dots, l-1, l$ ($2l+1$ numbers) gives the projection of the angular momentum operator along the z -direction or the expected value of \hat{L}_z . An analytic form is known for $\{Y_l^m(\theta, \phi)\}$, then the remaining problem is to find $\{u_{n,l}(r)\}$ that satisfies the radial equation given as

$$\left[-\frac{1}{2} \frac{d^2}{dr^2} + \frac{l(l+1)}{2r^2} + v_{eff}(r) \right] u_{n,l}(r) = \epsilon_{n,l} u_{n,l}(r). \quad (2.68)$$

Considering a hydrogen-like effective potential $v_{eff}(r) = -Z_{eff}/r$, an analytic form is known [52, 53] for $R_{n,l}(r) = u_{n,l}(r)/r$ which are the radial wavefunctions for the hydrogen atom. Nonetheless, this definition of radial functions is too complex to be adopted in a basis set construction and, therefore, not very useful. Instead, a simpler formulation is generally adopted, called Slater-type orbitals [54, 55], that incorporates important aspects such as the exponential decaying with r and the polynomial degree of $R_{n,l}(r)$. In Slater-type orbitals, $R_{n,l}(r)$ is defined as

$$R_{n,l}(r) = N_{n,l} r^{n-1} e^{-\zeta_{n,l} r}, \quad (2.69)$$

where $N_{n,l}$ is a normalization constant and the decaying rate $\zeta_{n,l}$ is related with the nuclear effective charge Z_{eff} (the nuclear charge Z partly screened by the electronic cloud) that can be estimated by Slater's rule [55].

Indeed, Slater-type orbitals are a good class of basis functions for molecular orbital expansion, they even have the right exponential decaying at large r values. Nonetheless, in HF and DFT calculations for polyatomic systems, two-electron integrals (Eq. 2.7) of basis functions centered at different atomic sites must be calculated, and this task can be very hard to

accomplish using Slater-type orbitals [54]. A great simplification is obtained by replacing the exponential decaying in Eq. 2.69 with a Gaussian-type decaying. This comes from the fact that the product of two Gaussian functions centered in different points A and B, can be rewritten as a third Gaussian centered in point a P in the line joining A and B. Then, the multicenter two-electron integrals can readily be solved. In Gaussian-type orbitals, the radial function $R_n(r)$ is defined as

$$R_n(r) = N_n r^{n-1} e^{-\alpha r^2}, \quad (2.70)$$

where N_n is a normalization constant, and α the decaying rate. An additional simplification adopted is using only atomic orbitals with $l = n - 1$ [54], *e.g.*, $1s$, $2p$, $3d$, $4f$, and so on, to represent all atomic states.

In general, Gaussian-type orbitals are used as a function of cartesian coordinates (x, y, z) , so that the integrals of basis functions are carried out separately for each coordinate. In this case, the Gaussian-type orbitals [56] are defined as

$$g(x, y, z; \alpha, i, j, k) = \left(\frac{2\alpha}{\pi}\right)^{3/4} \left[\frac{(8\alpha)^{i+j+k} i! j! k!}{(2i)!(2j)!(2k)!}\right]^{1/2} x^i y^j z^k e^{-\alpha(x^2+y^2+z^2)}, \quad (2.71)$$

where i , j , and k are integer numbers higher than or equal to zero, and they satisfy the relation $i + j + k = l$. It is important to notice that these functions present a higher degeneracy than canonical atomic orbitals for $l \geq 2$, the number of different combinations (i, j, k) given by $(l + 1)(l + 2)/2$ [55]. For example, if $l = 2$ there are 6 possible Gaussian orbitals while 5 canonical orbitals ($m = -2, -1, 0, 1, 2$). In fact, the 6 d -type Gaussian orbitals (d_{xy} , d_{yz} , d_{xz} , d_{x^2} , d_{y^2} , d_{z^2}) can be combined to form the 5 d -type canonical orbitals (d_{xy} , d_{yz} , d_{xz} , $d_{x^2-y^2}$, d_{z^2}).

In this work, we have adopted the Pople basis set 6-31G(d,p) [57, 54] for the study of isolated molecules in NWChem code [39]. This basis set adopts Contracted Gaussian Orbitals (CGOs) with 6 Gaussian orbitals for atomic core states and split basis functions for atomic valence states, *i.e.*, each atomic valence state has two basis functions: one CGO with 3 Gaussian orbitals and the most diffuse Gaussian orbital left uncontracted. For example, in the case of a second-row element in the p-block, the basis set would be defined as

$$\begin{aligned} \phi_{1s} &= \sum_{i=1}^6 d_{i,1s} g_{1s}(\alpha_{i,1s}) \\ \phi'_{2s} &= \sum_{i=1}^3 d'_{i,2s} g_{1s}(\alpha'_{i,2sp}) \\ \phi''_{2s} &= g_{1s}(\alpha''_{2sp}) \\ \phi'_{2p} &= \sum_{i=1}^3 d'_{i,2p} g_{2p}(\alpha'_{i,2sp}) \\ \phi''_{2p} &= g_{2p}(\alpha''_{2sp}), \end{aligned} \quad (2.72)$$

where the scheme adopted for the basis functions of the $2p$ -orbital must be applied for its three cartesian components $2p_x$, $2p_y$, and $2p_z$. The same α -exponent value is adopted for atomic

orbitals in the same shell, where the diffuse functions ϕ_{2s}'' and ϕ_{2p}'' share the same Gaussian exponent α_{2sp}'' , and the contracted functions ϕ_{2s}' and ϕ_{2p}' share the same set of Gaussian exponents $\{\alpha_{i,2sp}'\}$. Additionally, in the 6-31G(d,p) basis set, the d -type orbitals (uncontracted) are included for second and third-row elements and the p -type orbitals for the hydrogen atom. These additional functions are the so-called polarization functions, and they improve over the electronic cloud configuration on the molecular environment. All basis set coefficients and exponents are optimized to minimize the total energy value in SCF calculations.

Another important CGO basis set scheme, broadly adopted, was developed by Dunning and co-workers [58] to converge post-HF calculations (correlation corrections) systematically to the complete basis set limit. These basis sets are classified as cc-pVNZ, N=D,T,3,4,5,..., where ‘cc-p’ stands for ‘correlation-consistent polarized’ and ‘V’ indicates that they are converged for valence-only calculations. In this work, we used the cc-pVDZ (valence double-zeta) basis set at MP2 calculations performed in NWChem. To exemplify, for a second-row element in the p -block, we have the set of basis functions $1s$, $2s'$, $2s''$, $2p'$, $2p''$, and $3d$ as an additional polarization function (adding up to 14 basis functions). As discussed before, these basis functions are contractions of Gaussian orbitals, however, in this specific case, they adopt spherical harmonics for the angular part.

A different strategy adopted for basis set generation is using Numeric Atom-centered Orbitals (NAOs) instead of analytical orbitals as described before. These basis functions are implemented in the FHI-aims code [59] whose basis set generation procedure is described below. In this case, the radial equation (Eq. 2.68) is numerically solved, for a certain atomic specie, on a logarithmic radial grid defined as

$$r(i) = r_0 e^{(i-1)\alpha}, \quad (i = 1, \dots, N_{\log}), \quad (2.73)$$

and smooth radial functions $u_{n,l}(r)$ are then obtained through cubic spline interpolation. When a calculation for a multi-atomic system is launched in FHI-aims code, a minimal basis set is constructed for each required atomic species, through a non-spinpolarized SCF calculation using DFT-LDA or DFT-GGA, depending on the type of calculation addressed for the multi-atomic system. Additional tabulated basis functions are then included, necessary for the correct description of the wavefunction changes during bond formation in the material. These additional radial functions are obtained using different effective potentials $v_{eff}(r)$ (hydrogen-like with integer or fractional values of Z_{eff} , and doubly positive charged free ions), therefore, they are not necessarily orthonormal to one another. Accordingly, the Gram-Schmidt method is applied for each atomic species, explicitly on the logarithmic grid, to orthonormalize the radial functions at the same angular momentum channel l . Differently from the above-discussed analytical basis functions, the NAOs can intrinsically capture wavefunctions oscillations in the atomic core region and have the correct decaying profile at large r values. Furthermore, onset errors are avoided through the minimal basis sets generation process described above since these basis functions have been previously converged for each isolated atomic species.

A remarkable method adopted in the NAOs generation, for enhancing their computational efficiency, is including a cutoff potential ($v_{\text{cut}}(r)$) in $v_{\text{eff}}(r)$ (Eq. 2.68) to prevent slow-decaying radial functions tail that has a minor contribution in numerical integrations. In the FHI-aims code [59], $v_{\text{cut}}(r)$ is defined as

$$v_{\text{cut}}(r) = \begin{cases} 0 & r \leq r_{\text{onset}} \\ \left[\frac{s}{(r - r_{\text{cut}})^2} \right] e^{\left(\frac{w}{r_{\text{onset}} - r} \right)} & r_{\text{onset}} < r < r_{\text{cut}} \\ \infty & r \geq r_{\text{cut}} \end{cases}, \quad (2.74)$$

which starts from zero for r smaller than r_{onset} and grows smoothly to infinity during an interval $w = r_{\text{cut}} - r_{\text{onset}}$ that by default is set as 2.0Å. s is a scaling parameter and, by default, it is set at 200Ha. Since radial functions of different angular momentum l have different decaying behaviors, for example, d and f atomic orbitals are localized closer to the nucleus than s and p orbitals, a different value of $r_{\text{onset},i}$ is defined for each radial function $u_i(r)$ ($i = n, l$). Hence, $r_{\text{onset},i}$ is determined by

$$\int_{r_{\text{onset},i}}^{\infty} dr |u_i(r)|^2 = \eta_{\text{cut}}, \quad (2.75)$$

where η_{cut} is a small threshold in the order of 1E-4.

The additional basis functions commented on above are determined through a basis set generation scheme described in Ref. [59]. Basically, the average energy error of a set of N_d non-spinpolarized homodimers ($N_d \sim 4$ -5 different bond distances d_i between two atoms of the same chemical element) given by

$$\Delta_{\text{basis}} = \frac{1}{N_d} \sum_{i=1}^{N_d} \{E_{\text{basis}}(d_i) - E_{\text{cb}}(d_i)\}, \quad (2.76)$$

where the energies are obtained through non-SCF calculations using DFT-LDA, is minimized through basis set extrapolation. The reference energy values $\{E_{\text{cb}}(d_i)\}$ are obtained with a huge systematically converged basis set of atomic excited-state functions. The additional basis functions are included systematically by accepting, at each turn, only the function with the highest impact on Δ_{basis} . Interestingly, the selected sets of additional basis functions naturally group as tiers of increasing angular momentum functions, *e.g.*, spd - $spdfg$ - $spdf$ -... for oxygen. The inclusion of additional tiers increases the total energy accuracy. For light elements, such as the ones we are interested in this work, the ‘‘tier2’’, which corresponds to the inclusion of the additional basis functions of the first and second tiers, already provides good structural results.

In the work by Zhang *et al.* [60], they describe a procedure for the generation of tabulated NAOs with valence-correlation consistency (NAO-VCC- nZ), similar to the Dunning basis sets cc - $pVnZ$ [58], for correlation-correction calculations which involve unoccupied states.

2.5.2 Planewave basis set (PW)

In order to represent a periodic structure we must rely on a very important concept, the Bravais lattice. A good definition for the Bravais lattice is given by Ashcroft and Mermin [61] as an infinite array of points arranged so that it looks exactly the same from whichever points this array is looked at. Accordingly, a Bravais lattice is defined as a set of points written as

$$\vec{R} = n_1\vec{a}_1 + n_2\vec{a}_2 + n_3\vec{a}_3, \quad n_1, n_2, n_3 \in \mathbb{Z}, \quad (2.77)$$

where \vec{a}_1 , \vec{a}_2 , and \vec{a}_3 are translation vectors that are not all on the same plane. Any crystalline structure can then be represented by an atomic basis and a Bravais lattice to describe the positions where this set of atoms must be repeated to form the periodic structure. By convention, the cell delimited by the lattice vectors $\{\vec{a}_1, \vec{a}_2, \vec{a}_3\}$ is named a conventional unit cell.

In a crystal, the electrons will be subject to a periodic potential which obeys the relation $\hat{U}(\vec{r}) = \hat{U}(\vec{r} + \vec{R})$, where \vec{R} is of the form defined in Eq. 2.77. Accordingly, the electronic wavefunctions $\{\psi_i(\vec{r})\}$ must also have the periodicity of the lattice, *i.e.*, $\psi_i(\vec{r}) = \psi_i(\vec{r} + \vec{R})$. In this case, a convenient basis set for the expansion of the electronic wavefunctions $\{\psi_i(\vec{r})\}$ are the set of eigenstates $\{\varphi_{\vec{G}}(\vec{r})\}$ of a non-interacting electron gas under the periodic boundary conditions of this same lattice. Thus, each $\varphi_{\vec{G}}(\vec{r})$ function must also obey the relation $\varphi_{\vec{G}}(\vec{r}) = \varphi_{\vec{G}}(\vec{r} + \vec{R})$. These functions, called as planewaves, are defined as

$$\varphi_{\vec{G}}(\vec{r}) = \frac{e^{i\vec{G}\cdot\vec{r}}}{\sqrt{V}}, \quad (2.78)$$

where the set of \vec{G} -vectors define the so-called reciprocal lattice and are given as

$$\vec{G} = n_1\vec{b}_1 + n_2\vec{b}_2 + n_3\vec{b}_3, \quad n_1, n_2, n_3 \in \mathbb{Z}, \quad (2.79)$$

with \vec{b}_1 , \vec{b}_2 , and \vec{b}_3 the lattice vectors of the reciprocal lattice:

$$\begin{aligned} \vec{b}_1 &= \frac{2\pi}{V}(\vec{a}_2 \times \vec{a}_3), \\ \vec{b}_2 &= \frac{2\pi}{V}(\vec{a}_3 \times \vec{a}_1), \\ \vec{b}_3 &= \frac{2\pi}{V}(\vec{a}_1 \times \vec{a}_2), \end{aligned} \quad (2.80)$$

satisfying the relation $\vec{b}_i \cdot \vec{a}_j = 2\pi\delta_{ij}$, and $V = \vec{a}_1 \cdot (\vec{a}_2 \times \vec{a}_3)$ is the volume of the unit cell for the direct space. These states have a dependence on the \vec{G} -vectors which determine their energy values given by $\varepsilon_{\vec{G}} = \hbar^2\vec{G}^2/2m$. In general, concerning the use of planewaves as basis functions, the size of the basis set is determined through an energy cutoff for $\varepsilon_{\vec{G}}$ as the minimum acceptable value. The cutoff energy for the basis set will thus depend on the system under study since higher \vec{G} vectors are necessary to represent tiny oscillations near the atomic cores.

Concerning crystalline systems, Bloch's theorem [61] states that any electronic state due to a periodic potential must satisfy the following condition:

$$\psi_{n\vec{k}}(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}}\psi_{n\vec{k}}(\vec{r}), \quad (2.81)$$

where n stands for the band index and \vec{k} for the vectors lying inside a primitive unit cell of the reciprocal lattice, defined as

$$\vec{k} = \frac{m_1}{N_1}\vec{b}_1 + \frac{m_2}{N_2}\vec{b}_2 + \frac{m_3}{N_3}\vec{b}_3, \quad m_i = -N_i, -N_i + 1, -N_i + 2, \dots, N_i. \quad (2.82)$$

$N_1, N_2,$ and N_3 are the number of unit cells in the Bravais lattice repeated along each direction $\vec{a}_1, \vec{a}_2,$ and \vec{a}_3 . Although the definition in Eq. 2.82 is in terms of the conventional unit cell of the reciprocal lattice, in general, the \vec{k} -vectors are defined with respect to the more convenient Wigner-Seitz unit cell [61] or first Brillouin Zone (BZ). The BZ can be defined as the volume in the reciprocal space delimited by the planes that perpendicularly cross the lattice vectors of the primitive cell $\pm\vec{b}_1, \pm\vec{b}_2,$ and $\pm\vec{b}_3$ at half-length. As the BZ is also a primitive unit cell of the Reciprocal lattice, it contains all the information about the electronic structure of a crystalline system. It further displays all the symmetry elements contained in the electronic structure of the system. A convenient set of $\{\vec{k}\}$ vectors that exploit the high symmetry of the BZ is defined in the work by Monkhorst and Pack [62] and are given by

$$\vec{k} = u_1\vec{b}_1 + u_2\vec{b}_2 + u_3\vec{b}_3, \quad u_i = \frac{(2m_i - N_i - 1)}{2N_i}, \quad m_i = 1, 2, 3, \dots, N_i, \quad (2.83)$$

where $\{N_i\}$ are integer numbers that determine the number of \vec{k} -vectors in each direction $\vec{b}_1, \vec{b}_2,$ and \vec{b}_3 .

Since the electronic wavefunctions $\{\psi_{n\vec{k}}(\vec{r})\}$ have the periodicity of the Bravais lattice, a convenient basis set for their expansion are the planewaves defined in Eq. 2.78, as discussed before. On this basis set representation $\{\psi_{n\vec{k}}(\vec{r})\}$ will be written as

$$\psi_{n\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}} \sum_{\vec{G}} \psi_{n\vec{k}}(\vec{G}) e^{i\vec{G}\cdot\vec{r}}, \quad \psi_{n\vec{k}}(\vec{G}) = \frac{1}{\sqrt{V}} \int_V d\vec{r} e^{-i\vec{G}\cdot\vec{r}} \psi_{n\vec{k}}(\vec{r}), \quad (2.84)$$

where the set of \vec{G} -vectors to carry out the expansion will be determined by the cutoff energy described before. In this work, the electronic structure for crystalline systems was obtained using DFT. In this case, we must solve eigenvalue equations of the form given by

$$\hat{h}_{\text{KS}} \psi_{n\vec{k}}(\vec{r}) = \varepsilon_{n\vec{k}} \psi_{n\vec{k}}(\vec{r}), \quad (2.85)$$

where \hat{h}_{KS} is the Kohn-Sham Hamiltonian defined in Eq. 2.19. $\hat{h}_{\text{KS}}[n(\vec{r})]$ is a functional of the electronic charge density of the system that in the non-spinpolarized case can be defined as

$$n(\vec{r}) = 2 \sum_n \sum_{\vec{k}} f(\varepsilon_{n\vec{k}}) \psi_{n\vec{k}}^*(\vec{r}) \psi_{n\vec{k}}(\vec{r}). \quad (2.86)$$

$f(\varepsilon_{n\vec{k}})$ gives the occupation of the Kohn-Sham states and can be assumed as 1 or 0 for insulating materials or, for conducting systems, it will be given by

$$f(\varepsilon_{n\vec{k}}) = \int_{-\infty}^{\mu} d\varepsilon \tilde{\delta}(\varepsilon - \varepsilon_{n\vec{k}}), \quad (2.87)$$

where $\tilde{\delta}(\varepsilon')$ is a smearing function whose width will be determined by a broadening parameter σ , and μ is the chemical potential of the system [63]. The electronic charge density $n(\vec{r})$ will also be expanded in a planewave basis set, however, it must be adopted a higher energy cutoff than for the wavefunctions in this case since the product $\psi_{nk}^*(\vec{r})\psi_{nk}(\vec{r})$, given by

$$\psi_{nk}^*(\vec{r})\psi_{nk}(\vec{r}) = \frac{1}{V} \sum_{\vec{G}'} \sum_{\vec{G}} e^{i(\vec{G}-\vec{G}')\cdot\vec{r}} \psi_{nk}^*(\vec{G}')\psi_{nk}(\vec{G}), \quad (2.88)$$

presents expansion terms with wavevectors $\vec{G}'' = \vec{G} - \vec{G}'$ as large as $2\vec{G}$ ($\vec{G}' = -\vec{G}$). Since the energy cutoff of the plane waves basis set scales with \vec{G}^2 , we must adopt an energy cutoff four times bigger than the energy adopted for the expansion of the wavefunctions to have a good representation of the electronic charge density.

As commented before, electronic wavefunctions have an oscillatory behavior in the vicinity of atomic nuclei due to the singular character of the nuclear potential at the origin. These tiny oscillations are very hard to capture using a planewave expansion since it would be necessary to adopt an extremely large basis set, which is computationally prohibitive. However, in most cases, we are interested in physical phenomena depending on the electronic interactions far apart from the nuclei, where the electronic wavefunctions are sufficiently smooth to be represented in a planewave basis set, *e.g.*, chemical bonding, first ionization energies, crystalline band structure, optical excitations, etc. In this case, the valence atomic states are the most relevant for physical description, whereas the contribution of the core states is mainly to screen the nuclear potential felt by the valence electrons. In this regard, an important proposition was adopting effective atomic potentials that simultaneously account for the bare nuclear potential and the coulomb potential due to the core electrons. These effective potentials, called pseudopotentials, are constructed for the isolated atomic species and must provide atomic radial functions that are smooth near the origin and fit an all-electron solution for a radius higher than a certain cutoff radius.

There are many pseudopotential propositions in the literature [64, 65, 66, 67] as an attempt to address two essential points: computational efficiency and accuracy. The first point concerns the size of the planewave basis set required for the expansion of the pseudo wavefunctions, and the second is whether these wavefunctions and their respective charge densities can correctly describe the physical systems of interest. In this work, we have adopted Optimized Norm-Conserving Vanderbilt Pseudopotentials (ONCVPP) by Hamann [67]. It is an enhanced version of the Norm-Conserving Pseudopotential (NCPP) proposed by Hamann *et al.* [64], based on a proposition presented by Vanderbilt [65]. In general terms, the NCPP has the following properties described in Ref. [64]:

1. the all-electron and pseudo eigenvalues agree for a certain atomic configuration

$$\varepsilon^{\text{AE}} = \varepsilon^{\text{PP}} = \varepsilon; \quad (2.89)$$

2. the all-electron and pseudo radial wavefunctions agree beyond a cutoff radius r_c

$$R_l^{\text{AE}}(\varepsilon, r) = R_l^{\text{PP}}(\varepsilon, r), \quad r > r_c; \quad (2.90)$$

3. the integrals from 0 to r of the all-electron and the pseudo electronic charge densities agree for $r > r_c$ (norm conservation)

$$\int_0^r dr' r'^2 |R_l^{\text{AE}}(\varepsilon, r')|^2 = \int_0^r dr' r'^2 |R_l^{\text{PP}}(\varepsilon, r')|^2, \quad r > r_c; \quad (2.91)$$

4. the logarithmic derivatives of the all-electron and the pseudo radial wavefunctions and their respective first energy derivatives agree for $r > r_c$

$$\left. \frac{d}{dr'} \ln(R_l^{\text{AE}}(\varepsilon, r')) \right|_{r'=r} = \left. \frac{d}{dr'} \ln(R_l^{\text{PP}}(\varepsilon, r')) \right|_{r'=r} \quad (r > r_c) \quad (2.92)$$

$$\left. \frac{d}{d\varepsilon'} \frac{d}{dr'} \ln(R_l^{\text{AE}}(\varepsilon', r')) \right|_{(\varepsilon', r')=(\varepsilon, r)} = \left. \frac{d}{d\varepsilon'} \frac{d}{dr'} \ln(R_l^{\text{PP}}(\varepsilon', r')) \right|_{(\varepsilon', r')=(\varepsilon, r)}.$$

The general procedure for an NCPP generation is first solving the KS equations (Eq. 2.19) for an isolated atomic species to determine $R_l^{\text{AE}}(\varepsilon, r)$, using the relations presented above to generate $R_l^{\text{PP}}(\varepsilon, r)$ that are smooth near the nucleus, and inverting the KS-equations to obtain $\hat{V}_l^{\text{PP}}(\varepsilon, r)$ given by [55]

$$\hat{V}_l^{\text{PP}}(\varepsilon, r) = \varepsilon - \frac{l(l+1)}{2r^2} + \frac{1}{2rR_l^{\text{PP}}(\varepsilon, r)} \frac{d^2}{dr^2} [rR_l^{\text{PP}}(\varepsilon, r)]. \quad (2.93)$$

At the end, the Hartree and exchange-correlation potentials due to $R_l^{\text{PP}}(\varepsilon, r)$ are calculated and discounted from $\hat{V}_l^{\text{PP}}(\varepsilon, r)$ to obtain the bare-ion pseudopotential.

In fact, the pseudopotentials should be generated with the same kind of XC functional that will be applied to the system (molecule or crystal) electronic structure calculation. Regarding calculations with hybrid functionals, the generation of pseudopotentials in this category is still active [68] and continues to be explored. The common practice is enforcing the calculation with a hybrid functional even though the pseudopotentials had been generated with a GGA functional. In general, this leads only to tiny deviations from the all-electron results [68].

2.6 Computational packages

All the above-discussed methodologies and theories were employed through well-established computational packages in the materials science community, supported by active research groups in this field. In this section, we present a brief overview of these packages.

2.6.1 NWChem code

The NWChem code [39] is a quantum chemistry package freely distributed under the Open Source Educational Community Licence. It has the implementation of a variety of methodologies, *i.e.*, classical force field, HF, post-HF methods (including MP2), DFT, Time-Dependent DFT (TDDFT), and others, thus it can treat a large range of systems, containing from a few to a thousand atoms.

Here, we adopted the DFT and MP2 implementations of NWChem to study the structural properties of isolated molecules. As commented in Section 2.5.1, in this case, a local basis formulation is available, which is implemented to work with CGOs basis sets. Several basis set libraries are available in the NWChem distribution. Among them, in our work, we adopted Pople basis sets [57] 6-31g and the polarized version 6-31g(d,p), and the Dunning style basis set [58] cc-pVDZ. There is also the Basis Set Exchange repository [69] where many basis set definitions are available for the NWChem code. It is important to mention that by default all electrons are considered in the SCF calculations, which was adopted in this work.

An important capability available in NWChem is the possibility of defining internal coordinates to be used in geometry optimizations, *i.e.*, internal angles, bond distances, and torsion angles. This functionality allowed us to construct torsion potentials of dimer molecules by constraining their inter-ring torsion angles. Beyond that, NWChem has a helpful automatic detection tool for point group symmetries, or if necessary, they can also be specified manually. Regarding the optimization methods, we adopted the DRIVER module which has the implementation of a quasi-Newton optimization algorithm similar to BFGS. In this module, the convergence is achieved through the minimization of the Hellmann-Feymann forces (maximum component and root mean square of all components) with respect to the coordinates being used (cartesian, internals, ...), and the geometry modification in cartesian coordinates. We adopted the DEFAULT and the TIGHT convergence criteria in our calculations, defined in the NWChem manual [70].

2.6.2 FHI-aims code

The FHI-aims code [59] is a computer program for materials science simulations based on quantum-mechanical first principles. The fundamental methodology in FHI-aims is the DFT (also hybrid DFT), from where we can obtain the ground state total energy and derived quantities for molecules and solid-state materials. Additionally, it can be applied in calculations of single-quasiparticle excitation energies through the GW method and optical/neutral excitations (only available for isolated systems) through the Bethe-Salpeter Equation (BSE). There are other methodologies implemented in the FHI-aims code, such as perturbative corrections from RPA and MP2, but they were not explored in this work.

An important functionality in FHI-aims is its local basis implementation. It works with

basis sets of NAOs (see Section 2.5.1), carefully constructed by the team of developers to be highly efficient, accurate under basis set size convergence, and able to treat a wide range of systems. There are many tabulated basis sets available in the FHI-aims distribution, among them, the basis sets labeled “tier” perform very well in the description of structural and ground state electronic structure properties [59]. For each tier-type basis set, we can further improve upon the numerical integrations by adopting tighter integration grid settings, where the options are LIGHT, TIGHT, and REALLY_TIGHT, which control the fineness of the grid. In this work, we adopted tier2-TIGHT for obtaining the structural properties of the isolated molecules and the NAO-VCC-3Z [46, 60], *i.e.*, a correlation-consistent basis set, for the electronic structure properties, including the GW excitation energies and the optical excitations obtained through BSE. It is also important to highlight that FHI-aims is an all-electron method, however, its efficient numerical framework makes it computationally competitive with planewave-based platforms.

2.6.3 Quantum Espresso code (QE)

The QE code [71, 72] is a package for materials science simulation, based on a planewave formulation for the DFT. It is an open-source project, freely available to the entire community. There are several methodologies implemented for the calculation of the ground-state electronic structure of a broad range of systems, and further processing of these data into human-readable results. It is a modular package, thus many different codes are available in the QE distribution, each one with an independent input file structure.

The main module in QE is the pw.x, which is responsible for the calculation of the DFT ground state electronic structure using a planewave-based methodology (Section 2.5.2). It can further perform geometry relaxations and non-SCF calculations at specified k -points, this last only available for non-hybrid functionals. Many settings are required for running a calculation: type of calculation (relax, SCF, non-SCF, ...), ecutwfc (ecutrho), *i.e.*, energy cutoff for the planewave expansion of the wavefunctions (electronic density), Hamiltonian diagonalization method (the main are Davidson and the conjugated gradient), pseudopotential files, k -points mesh in the Brillouin-Zone (BZ), etc. These are some of the main settings, however, it depends on the kind of system addressed. For example, in this work, we included a Hubbard potential correction for the Zn 3d-states. We further adopted the hybrid-functional PBE0, which requires the additional specification of an auxiliary q -points mesh for evaluating the Hartree-Fock exchange.

Other important modules were used to process the ground state data, namely average.x to get the average electrostatic potential along a direction, pp.x to obtain orbitals electronic charge density, band.x to construct band structures, and projwfc.x to evaluate the projected density of states (PDOS) on the atomic orbitals.

Thiophene-based oligomers

Thiophene is a fundamental chemical compound in organic electronics. Among its several different formulations, thiophene linear oligomers, or oligothiophenes, are an important class of π -conjugated materials due to their broad range of favorable characteristics [2], *e.g.*, good charge mobility, sunlight harvesting in the UV-Vis range, good stability at ambient conditions, consolidated synthetic routes, and more. In this chapter, we present a theoretical investigation of specific thiophene-based systems focusing on their application in the active layer of organic solar cells.

The first class of systems investigated here are thiophene-furan co-oligomers. These materials have been proposed to address some issues concerning oligothiophenes application, namely low solubility rates in common organic solvents, and the non-planarity of the linear chains [5, 6]. Although furan oligomers (oligofurans) do not pose such problems, they exhibit low stability at ambient conditions. Then, an interesting attempt is the co-polymerization of thiophene and furan to provide optimized material properties. In this regard, we investigate the role of chain composition and size on the structural and optoelectronic properties of thiophene-furan co-oligomers with up to 4 units. We present a comparative study of these systems with thiophene and furan-only oligomers.

In another direction, the design of oligomers with special ending groups is very relevant for device building. Specifically, here we focus on cyanoacrylate end-capped oligothiophenes (*n*TCs), which are expected to display a photoexcited donor-acceptor character as their Highest-Occupied Molecular Orbital (HOMO) is localized on the oligothiophene moiety and Lowest-Unoccupied Molecular Orbital (LUMO) on the cyanoacrylate termination [18, 21]. Such systems can be applied in the active layer of DSSCs [19, 24] since an effective electron-hole charge separation leads to better device performance. We present an investigation of the structural and optoelectronic properties of *n*TCs with 1 and 3 thiophene units.

Regarding the thiophene and furan oligomers/co-oligomers addressed here, the structural properties were studied with DFT and MP2 in NWChem. The T and F monomers were relaxed with constrained C_{2v} point-group symmetry using PBE0/6-31g(d,p). The optimizations were performed through the DRIVER module with the DEFAULT set of convergence criteria and SCF convergence energy of 1E-6 Ha. Unless otherwise mentioned, this setup was adopted for

all geometry optimizations performed with NWChem. We adopted the optimized geometries of the monomers as the starting point for the dimer geometries. For each dimer species, *i.e.*, 2T (Figure 3.1), 2F, and TF, we performed geometry optimizations using PBE0/cc-pVDZ, with constrained $X-C-C-X$ dihedral angles (X being S or O) from 0° to 180° , at 10° steps, and constrained point-group symmetry when applicable. For the optimized geometries, we calculated the total energies using MP2/cc-pVDZ to build the torsion potentials of the dimers. The dimer geometries with dihedral angle values of 0° , 40° , 90° , 150° , and 180° were also relaxed using MP2/cc-pVDZ, with the same constraints as before. We adopted the most stable geometries of the dimers as the starting point for the trimer and tetramer geometries, which were then relaxed using PBE0/6-31g(d,p), with constrained point-group symmetry when applicable. For the trimers and tetramers containing TT sequences, the relaxations were carried out with constrained $S-C-C-S$ dihedral angles at 150° . The most stable geometries of 2T, 3T, 4T, and FTTF were further relaxed, with no dihedral-angle constraints, using PBE0/cc-pVDZ and MP2/cc-pVDZ for comparison.

For all systems addressed in this chapter, the electronic structure results were obtained using PBE0/NAO-VCC-3Z in FHI-aims, with SCF convergence energy of $1E-6eV$. Starting from the PBE0 ground state, we calculated the ionization energy spectrum with the GW method (G0W0 approximation). Regarding the integration of the correlation part of the self-energy along the imaginary frequency axis, the AC technique was adopted through a two-pole model function fitted at 40 frequency points in a modified Gauss-Legendre grid. The BSE was solved on top of the G0W0/PBE0 electronic structure to obtain the absorption spectra.

3.1 Thiophene linear oligomers

Thiophene (T) is an aromatic compound with a cyclic structure containing four π -conjugated carbon atoms and one sulfur heteroatom with sp^3 hybridization. It can be polymerized into stable π -conjugated linear chains nT ($n = 2, 3, 4, \dots$) displaying suitable optoelectronic properties for device application.

The geometry of the shortest linear chain 2T is shown in Figure 3.1. To investigate the 2T conformation, which is the basis for understanding more extended chains, we calculated the 2T torsion-potential curve using PBE0 and MP2, a standard benchmark method for such calculations [16, 73, 74]. In Figure 3.1-a we present the PBE0 and MP2 results. As we see, both methods agree in the torsion-potential overall profile, although providing different energy-barrier heights at 0° , 90° , and 180° , and minima at slightly different angular positions. The stationary points of the 2T torsion potential are also presented in Table 3.2. The *cis*-torsioned minimum is obtained around 30° with PBE0 and around 40° with MP2. The methods also predict the 2T *trans*-torsioned minimum at different positions, namely around 160° with PBE0 and 150° with MP2. These differences have already been reported in other theoretical works

[73, 74], while the most accepted experimental values are 36° and 148° for the *cis*-torsioned and *trans*-torsioned minima, respectively [75]. Bloom and Wheeler [74] adopted a methodology called Focal Point Analyses (FPA) to investigate the torsion potentials of bithiophene and bifuran, here called 2T and 2F. In FPA, a simultaneous convergence of the basis set size is performed, there [74] a Dunning basis set cc-pVXZ ($X=D, T, Q,$ and 5), and of successive inclusion of post-HF corrections (correlation effects) going through MP2, CCSD, CCSD(T) and CCSDT. By extrapolation, they obtained the 2T torsion-potential minima at 32.3° and 156.3° , in better overall agreement with our MP2 results. Regarding the energy-barrier height at 180° , our PBE0 result of 0.005eV agrees with FPA (around 0.005eV) and the experimental value of about 0.003eV . For the transition state at 90° , the energy-barrier height with MP2 is in good accordance with FPA, while the energy value with PBE0 is overestimated. Thus, although differences were seen between the PBE0 and MP2 results, both methods indicate that 2T will more probably assume a *trans*-torsioned conformation, even if at ambient conditions $k_B T = 0.026\text{eV}$. Then, we proceeded by adopting the 2T *trans*-torsioned conformer with an *S-C-C-S* dihedral angle of 150° .

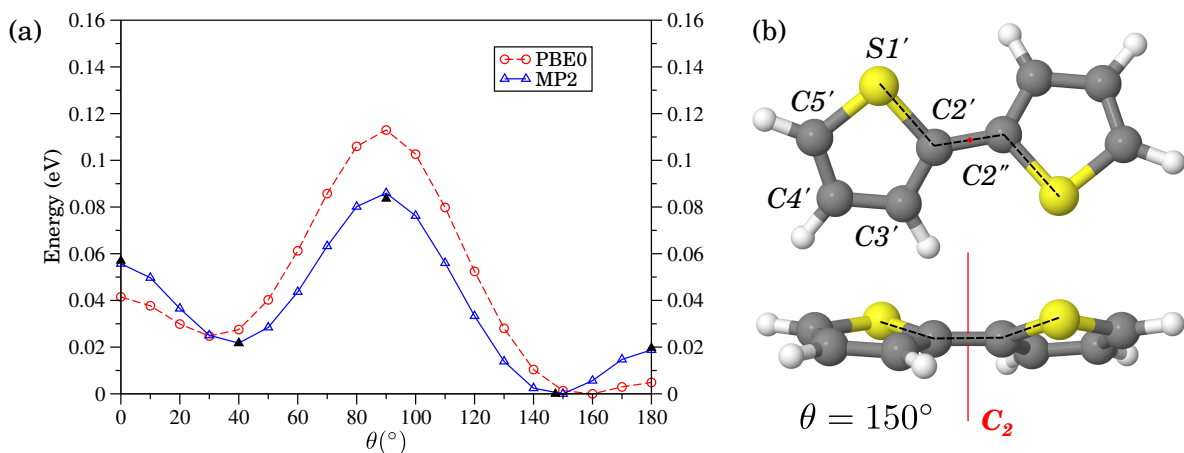


Figure 3.1: (a) 2T torsion potential calculated using PBE0 (circles) and MP2 on top of PBE0 geometries (open triangles). The solid triangles concern the energy values obtained with MP2 relaxations. For each method, energies are aligned to the respective global minimum. (b) Top and side views of 2T geometry: the dihedral angle θ is indicated by the black dashed line ($\theta=150^\circ$ in the corresponding figure).

In Table 3.1 we present the bond lengths and bond angles of the 2T *trans*-torsioned conformer obtained with PBE0 and MP2. Since 2T has a C_2 -symmetry axis crossing its $C2'-C2''$ bond, perpendicular to the plane of the page in Figure 3.1-b, the results presented in Table 3.1 can be extended to all the parameters which are equivalent by symmetry. As we see, the PBE0 and MP2 geometries are very similar and there is a good agreement between the theoretical results and the experimental values by gas-phase electron-diffraction technique. These measurements were performed for a 2T gas at a temperature of around 97.5° and, as commented

before, a *cis*-torsioned and a *trans*-torsioned conformation was found with dihedral-angle values of 36° and 148° , respectively, and populations of 44% and 56% [75, 76].

Table 3.1: Bond lengths (\AA) and bond angles ($^\circ$) of the 2T *trans*-torsioned conformer obtained through PBE0 and MP2 relaxations with the dihedral angle constrained to $\theta = 150^\circ$. Experimental data from electron and X-ray diffraction techniques taken from Ref. [76].

	PBE0	MP2	electron diffraction	X-ray diffraction
$S1' - C2'$	1.741	1.739	1.733	1.713
$C2' - C3'$	1.377	1.393	1.370	1.432
$C3' - C4'$	1.420	1.419	1.452	1.444
$C4' - C5'$	1.368	1.386	1.363	1.357
$C5' - S1'$	1.723	1.724	1.719	1.698
$C2' - C2''$	1.451	1.455	1.456	1.448
$S1' - C2' - C3'$	110.4	110.6	111.8	112.5
$C2' - C3' - C4'$	113.3	113.0	111.9	108.0
$C3' - C4' - C5'$	112.7	112.6	112.3	114.9
$C4' - C5' - S1'$	111.6	111.6	112.3	112.1
$C5' - S1' - C2'$	92.0	92.2	91.5	92.5
$S1' - C2' - C2''$	120.6	120.6	121.9	121.2

The 3T and 4T initial geometries, shown in Figure 3.2, were based on 2T: the 3T with a σ_s -symmetry plane crossing its central ring (perpendicular to the plane of the page in the top geometry of Figure 3.2-a), passing through the $C3''$ - $C4''$ bond and the $S1''$ atom, and the 4T with a C_2 -symmetry axis (perpendicular to the plane of the page in the top geometry of Figure 3.2-b) crossing the $C5''$ - $C5'''$ bond. These initial geometries were then relaxed using PBE0/6-31g(d,p) with constrained *S-C-C-S* dihedral-angle values at 150° . We further relaxed these geometries using PBE0/cc-pVDZ and MP2/cc-pVDZ, with no dihedral constraints, to understand the effect of chain size on the *S-C-C-S* dihedral angles.

In Figure 3.2 we show the 3T and 4T final geometries. As we see, they present a slightly curved structure. Regarding the *S-C-C-S* dihedral-angle values obtained through PBE0/cc-pVDZ and MP2/cc-pVDZ relaxations, the results are consistent with those observed for 2T, *i.e.*, around 160° with PBE0 and 150° with MP2, with only a small increase with increasing chain size in both cases. The bond lengths and bond angles of the side rings are also quite similar to those obtained for 2T, as shown in Table A.1, whereas the central rings are a little more stretched along the chain direction, resulting in a very small deviation concerning bond angles.

Moving to the electronic structure properties of oligothiophenes, in Figure 3.3 we present

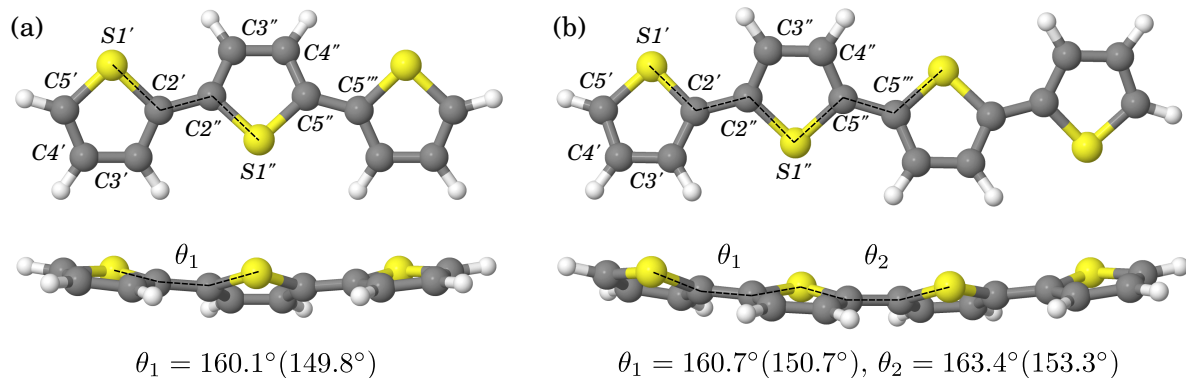


Figure 3.2: Top and side views of 3T (a) and 4T (b) geometries obtained using PBE0/6-31g(d,p) with constrained $S-C-C-S$ dihedral-angle values at 150° . The presented dihedral-angle values were obtained through further relaxations using PBE0/cc-pVDZ (MP2/cc-pVDZ) with no dihedral constraints.

the GW HOMO and LUMO energy levels for the thiophene systems discussed above, with 1 to 4 thiophene rings. We see from the electronic wavefunctions that the oligothiophenes HOMO and LUMO have a π -type character and are delocalized over the entire molecule. The HOMO energy increases with increasing chain size and the LUMO energy decreases, closing the HOMO-LUMO gap until saturation is reached, as evinced by the downtrend in the gap closure. Furthermore, an important characteristic of oligothiophenes concerning device application is the appropriate alignment of the HOMO and LUMO energy levels with consolidated semiconductor materials such as TiO_2 and ZnO . Regarding photovoltaic applications, the HOMO energy must be placed inside the gap of the semiconductor, and the LUMO energy in the conduction band (CB), as is verified for the oligothiophenes. Furthermore, as commented before, oligothiophenes have good stability at ambient conditions, a property that is attributed to their deeper HOMO energy level in comparison with other π -conjugated systems suitable for device application [2, 6].

In Figure 3.4 we present the optical excitations calculated for the oligothiophenes with BSE. We indicate with blue arrows the maximum position of the experimental absorption peaks reported in the literature, and for T and 2T, for which we have experimental data for the two first absorption peaks, the relative peak intensities are indicated by blue horizontal bars. As we see, there is good agreement between the results obtained in this work and the experimental data. Furthermore, our results agree with other theoretical works employing post-Hartree-Fock methods [79] and quasiparticle-self-consistent GW [80]. In the work by Prlj *et al.* [79], they show an inconsistency between TDDFT results for the first absorption peak of T and more accurate post-HF methods. In general, TDDFT presents the first absorption peak characterized by two $\pi - \pi^*$ bright transitions: the lowest-energy transition, mainly HOMO \rightarrow LUMO, and the highest-energy transition, mainly HOMO-1 \rightarrow LUMO. The opposite order is obtained with post-HF methods, and also here, using BSE/GW, where the brightest transition is mainly

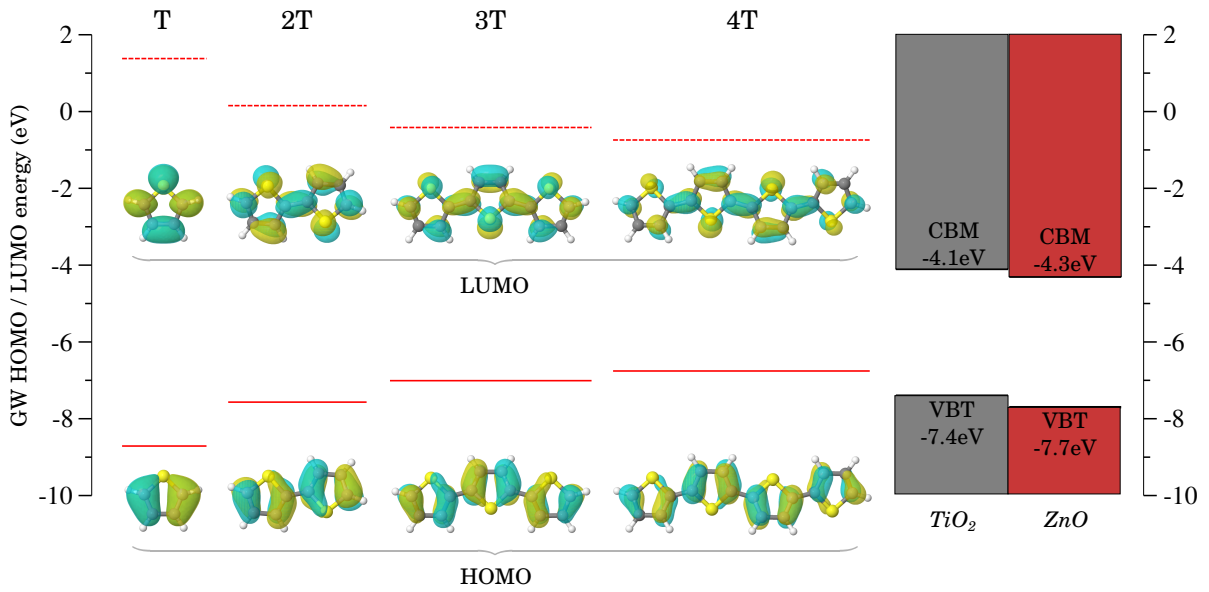


Figure 3.3: GW HOMO and LUMO energy levels for T, 2T, 3T, and 4T. The reference values for the CBM and VBT of TiO_2 and ZnO structures are relative to the vacuum level and were obtained from experimental data in Ref. [77, 78]. Wavefunctions isosurface cutoff of 0.08.

HOMO→LUMO (90%).

For the chains with 2 to 4 thiophene rings, the first peak is characterized by only one $\pi - \pi^*$ bright transition, mainly HOMO→LUMO (more than 90%). The intensity of this HOMO→LUMO transition increases with increasing chain size, becoming the brightest transition already in the 2T. Since, as discussed before, the HOMO and LUMO orbitals are delocalized over the entire molecule, the transition dipole moment (Eq. 2.59) for this HOMO→LUMO transition, directed along the chain direction, increases with increasing chain size resulting in the rise of the oscillator strength f (f will stabilize at a certain chain size as indicated by the downtrend in its increment). Regarding the energy of this HOMO→LUMO transition, as shown in Figure 3.3, the HOMO-LUMO gap closes with increasing chain size and, consequently, the optical gap dictated by this transition. As presented in Table 3.5, it moves from around 5.5eV for the monomer to 3.1eV for the tetramer, in a better range for sunlight capture. This possibility of tailoring the optoelectronic properties of oligothiophenes makes these systems very versatile and applicable to a broad range of photovoltaic devices.

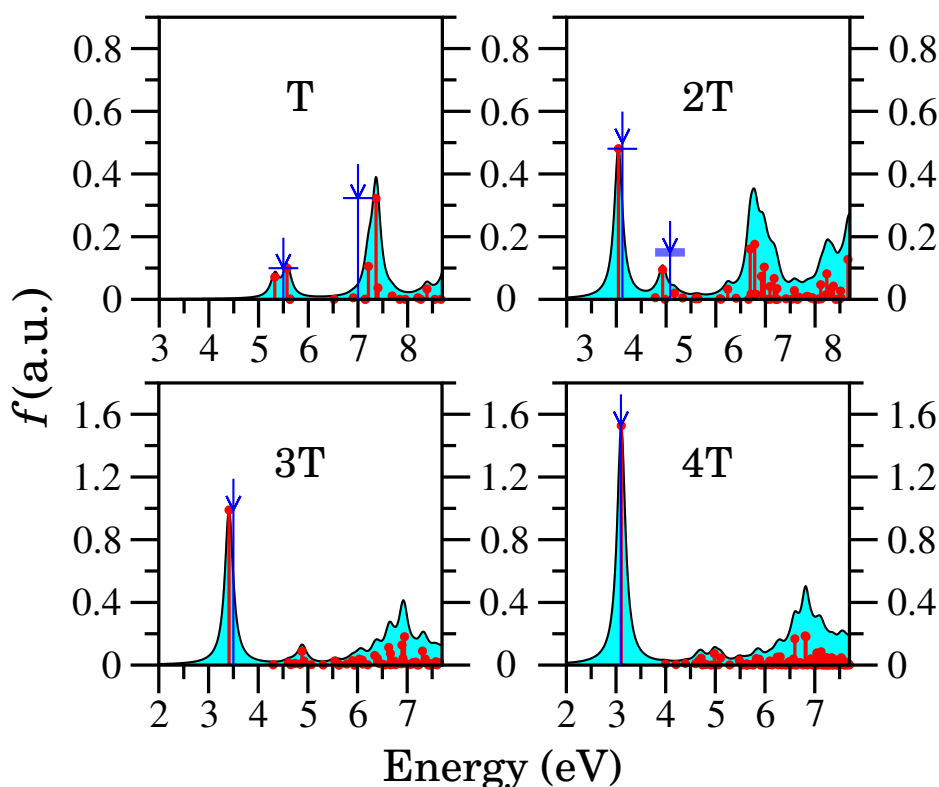


Figure 3.4: Absorption spectra for oligothiophenes with 1 to 4 thiophene units calculated with BSE/G0W0/PBE0. The red dots concern the excitation energies and the red bars are the respective oscillator strengths (f) given in arbitrary units (a.u.). The blue solid curves are an estimate of the absorption spectra of the molecules using Lorentzian functions with a scale parameter of 0.1 eV. The blue arrows indicate the position of the experimental absorption peaks (see Table 3.5) and the blue horizontal bars are an estimate of the relative peak intensities [81, 82, 83].

3.2 Thiophene-furan co-oligomers

Furan is a heterocyclic compound similar to thiophene but with an oxygen heteroatom instead of sulfur. It is a “green” electronic material, and its derived oligomers, *i.e.*, oligofurans, are expected to be biodegradable [11, 12]. Oligofurans show electronic properties similar to oligothiophenes [10, 84] and higher fluorescence quantum yields as reported in the literature [3]. They exhibit high solubility rates in common organic solvents, eliminating the necessity of chain alkylation for processing [5], and have a planar geometry, more rigid-to-torsion than oligothiophenes [10, 11, 74, 84]. Gidron *et al.* [3] show that the field effect mobility of thin films of oligofurans is similar to that of corresponding oligothiophene systems. On the other hand, oligofurans suffer from low stability in the presence of oxygen and light which can be harmful to the device lifetime [3].

Here, we present a theoretical investigation of the structural and optoelectronic properties

of thiophene-furan co-oligomers (with up to 4 rings), which are an alternative class of π -conjugated linear chains [6, 8, 9]. We compare their properties with those obtained for analog thiophene and furan-only oligomers. Regarding oligofurans classification, we adopt the same nomenclature scheme as that adopted for oligothiophenes in Section 3.1, *i.e.*, nF ($n = 2, 3, 4, \dots$), where F stands for Furan. For thiophene-furan co-oligomers, we follow chain ordering, *e.g.*, TFFT is a chain with two central F units and one end T unit at each side.

In Figure 3.5 we present the torsion potential curves obtained for 2T, 2F, and TF with PBE0 and MP2. As discussed in Section 3.1 for 2T, also here, for TF and 2F, the methods agree in the overall profiles. However, concerning 2F, we see a slightly different behavior around 0° : while the PBE0 curve rises monotonically for $O-C-C-O$ dihedral angles higher than 0° , the MP2 curve presents an almost flat profile up to 30° . In Table 3.2 we present the stationary points of the torsion potentials of the dimers. As indicated, the MP2 results for 2F predict a *cis*-torsioned conformation slightly lower in energy than the *cis*-planar, at a dihedral angle of around 10° . Bloom and Wheeler [74] obtained a similar result through FPA, with the *cis*-torsioned conformation at 16.3° and an energy value of 0.075eV about the global-minimum. The MP2 energy for the transition state of 2F is also in good agreement with the FPA results [74] of 0.173eV, whereas PBE0 predicts a higher energy value of 0.243eV. Regarding TF, our results are in good agreement with other theoretical works [16, 85]. Interestingly, it displays *cis*-planar and *trans*-planar stable conformations with an energy difference of only 0.012eV ($\approx 0.39k_B T$), differently from 2T and 2F.

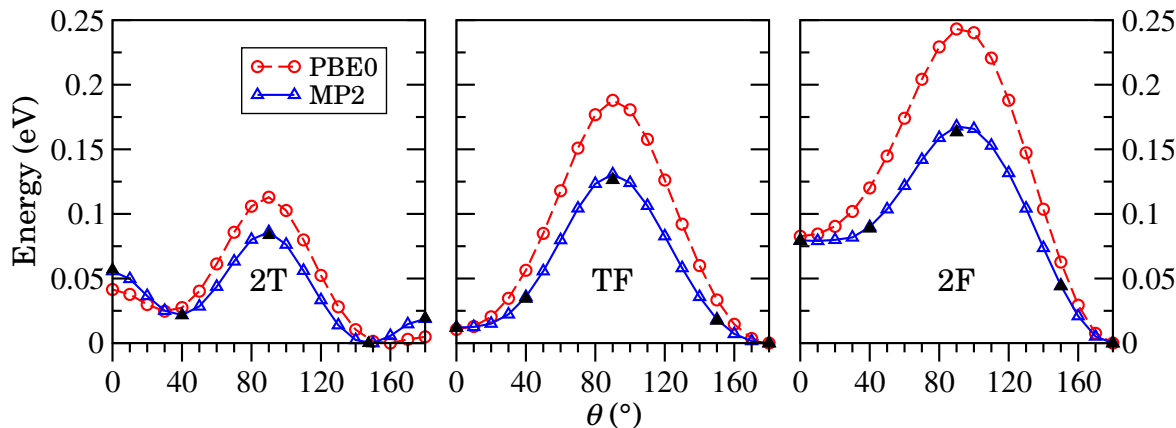


Figure 3.5: Torsion potential for 2T, TF, and 2F calculated using PBE0 (circles) and MP2 on top of PBE0 geometries (open triangles). The solid triangles concern the energy values obtained with MP2 relaxations. For each method, energies are evaluated relative to the respective global minimum energy.

In the following discussions, the results were obtained taking the most-stable conformers of the dimers, *i.e.*, at the global minimum of the torsion potentials shown in Figure 3.5. In Figure 3.6 we present the Molecular Electrostatic Potential (MEP) of 2T, TF, and 2F evaluated on their

Table 3.2: Stationary points of the torsion potential of 2T, 2F, and TF presented in Figure 3.5. The numbers out of parenthesis (in parenthesis) concern the values obtained with PBE0 (MP2): the first is θ in degrees, and the second is the energy in eV about the global minimum of the corresponding system.

	2T	TF	2F
<i>cis</i> -planar	0; 0.041 (0; 0.056)	0; 0.011 (0; 0.012)	0; 0.083 (0; 0.079)
<i>cis</i> -torsioned	30; 0.025 (40; 0.022)	–	– (10; 0.079)*
transition state	90; 0.113 (90; 0.086)	90; 0.188 (90; 0.131)	90; 0.243 (90; 0.168)
<i>trans</i> -torsioned	160; 0 (150; 0)	–	–
<i>trans</i> -planar	180; 0.005 (180; 0.019)	180; 0 (180; 0)	180; 0 (180; 0)

* MP2 energy slightly lower than that for the *cis*-planar conformation ($\approx 0.0004\text{eV}$).

vdW surface. It was calculated with the Jmol code [86] using the atomic Hirshfeld charges Q_i [87] obtained with PBE0, as shown in Table 3.3. Basically, the MEP at a point \vec{r}_j is the sum of all electrostatic contributions of the form $Q_i/|\vec{r}_j - \vec{r}_i|$, due to the atomic charges Q_i placed at \vec{r}_i . Accordingly, the MEP maps the charge distribution below the vdW surface, assuming a positive (negative) value over a positively (negatively) charged region. The MEP plots in Figure 3.6 indicate positive effective charge on sulfur and hydrogen atoms and negative on oxygen and carbon atoms. This is in good agreement with the Hirshfeld atomic charge values, except for the furan moieties, whose first carbon neighbors of the oxygen atom are positively charged; however, their effect on the MEP is suppressed by the large negative charge of the oxygen.

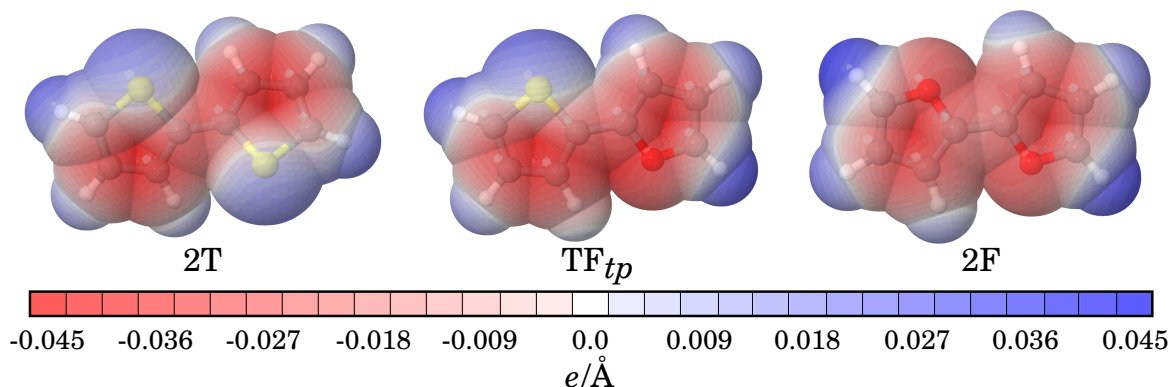


Figure 3.6: Molecular Electrostatic Potential (MEP) surface for 2T, TF_{tp} , and 2F. The MEP was evaluated on the vdW surface of the dimers.

The behavior of the torsion potentials in Figure 3.5 can be understood by the interplay between charge localization and the electrostatic interactions. The higher electronic confinement on the rings, brought on in non-planar conformers, leads to an increase in the electronic energy, as can be seen for the rise in the energy of the transition states at 90° . On the other hand, in a planar conformation, the effective electrostatic interaction between the atoms in the opposite

Table 3.3: Hirshfeld charges (e) of 2T, 2F, and TF calculated with PBE0. The atomic labels follow the scheme adopted in Figure 3.2 for the first two rings, where $X=S$ or O .

	2T	TF	2F
$X1'/X1''$	0.10	0.10/-0.07	-0.07
$C2'/C2''$	-0.04	-0.04/0.05	0.05
$C3'/C3''$	-0.07	-0.07/-0.09	-0.08
$H3'/H3''$	0.05	0.05	0.06
$C4'/C4''$	-0.07	-0.07/-0.08	-0.08
$H4'/H4''$	0.06	0.06	0.06
$C5'/C5''$	-0.08	-0.08/0.01	0.01
$H5'/H5''$	0.06	0.06	0.06

rings increases. For example, the *cys*-planar conformer of 2T (2F) experiences a repulsive interaction due to the S (O) atoms of the same polarity in the opposite rings and also the opposite hydrogen atoms. Regarding the TF conformer, there is an attractive interaction between S and O , with opposite charges, and a repulsive interaction between the hydrogen atoms. In the *trans*-planar conformation, 2T experiences repulsive interactions between S and H , TF experiences repulsive interaction between S and H and attractive interaction between O and H , and 2F only attractive interactions between O and H . The TF is the only dimer that experiences an attractive interaction in both planar conformations, namely between the opposite S and O atoms in the *cys*-planar and O and H in the *trans*-planar, allowing the rings to arrange in the plane of the molecule so that to move apart the charges of the same polarity.

Regarding the thiophene-furan co-oligomers with 3 and 4 rings, we obtain that avoiding a TT sequencing in the structure of the chains leads them to planarity, *i.e.*, TFT, FTF, TFTF, and TFFT have planar geometries, as expected from the above discussion. On the other hand, the FTTF presents a non-planar geometry with $O-C-C-S$ and $S-C-C-S$ dihedral angle values of 178.8° (177.1°) and 166.5° (153.8°) obtained through symmetry-constrained PBE0 (MP2) relaxations. Although, in this case, we performed geometry optimizations adopting only starting-point geometries with a *trans* configuration of the rings, or antiparallel, a similar result is expected in a *cys* or parallel configuration due to the symmetric character of the TF torsion potential (Figure 3.5). These conformational properties are relevant since they influence molecular packing, and consequently, π -stacking in solid-state materials. Accordingly, TT sequencing should be avoided to increase the charge mobility in thiophene-furan co-oligomers. As the polymerization mechanisms become more precise, this prescription can be followed to design materials with enhanced properties [8].

In Table 3.4 we present the GW ionization potential (IP_{GW}), electron affinity (EA_{GW}), and electronic gap (E_{gGW}) of the thiophene and furan linear chains. For the TF dimer, since the *cis*-planar (*cp*) and *trans*-planar (*tp*) conformers have a small energy difference of $\approx 0.39k_B T$,

Table 3.4: GW ionization potential (IP_{GW}), electron affinity (EA_{GW}), and electronic gap (E_{gGW}) of the thiophene and furan oligomers and co-oligomers. The Kohn–Sham IP values (IP_{KS}) obtained with PBE0 are also included, and the experimental data (IP_{exp}) is included when available.

	IP_{KS}	IP_{GW}	IP_{exp}	EA_{GW}	E_{gGW}
T	6.92	8.71	8.85 [88]	-1.38	10.09
F	6.68	8.68	8.87 [89]	-2.14	10.82
2T	6.10	7.57	7.83 [90]	-0.15	7.72
TF_{tp}	5.88	7.41	7.85 [91]	-0.32	7.73
TF_{cp}	5.86	7.36	7.85 [91]	-0.33	7.70
2F	5.78	7.40	7.90 [91]	-0.72	8.12
3T	5.74	7.01	7.38 [2]	0.41	6.59
TFT	5.51	6.84	-	0.27	6.57
FTF	5.50	6.89	-	0.28	6.61
3F	5.41	6.85	-	-0.10	6.95
4T	5.56	6.76	-	0.74	6.02
TFTF	5.31	6.54	-	0.61	5.94
TFFT	5.28	6.53	-	0.47	6.05
4F	5.22	6.53	-	0.25	6.28

probably displaying a similar population at ambient temperature, we present the results for both conformers. The IP_{GW} and EA_{GW} are given as minus the GW HOMO and GW LUMO energies, respectively, and the $E_{gGW}=IP_{GW}-EA_{GW}$. As a reference, we also present the Kohn–Sham IP (IP_{KS}) obtained with bare PBE0, and the experimental data by Photoelectron Spectroscopy (IP_{exp}) [2, 88, 89, 90, 91].

The one-shot GW approach on top of the PBE0 ground state provides IP values in good agreement with the experimental data. While our theoretical results predict slightly lower IP values for the furan-containing systems than for the oligothiophenes (with the same chain size), the experimental data predict similar IP values for all systems. Indeed, we see that the IP values have just a small dependence on the heteroatoms since the HOMO electronic density, shown in Figure 3.7, is mainly localized on the carbon atoms, differently from what is observed for the EA values. Regarding thiophene-furan co-oligomers, we see that the IP values are more strongly dictated by furan than by thiophene through their similarity with the IP values of the oligofurans. As observed for the oligothiophenes, also for oligofurans and thiophene-furan co-oligomers, the IP value decreases with the increasing number of chain moieties in agreement with the experimental data, *i.e.*, the HOMO state becomes less bonded with increasing chain size until a certain saturation. In contrast, the EA values increase from monomers to tetramers, leading to the electronic gap closure seen in Table 3.4. Interestingly, contrary to the IP values,

the electronic gap of the thiophene-furan co-oligomers is more strongly dictated by thiophene than by furan. We will show that this same behavior is observed for the optical gaps.

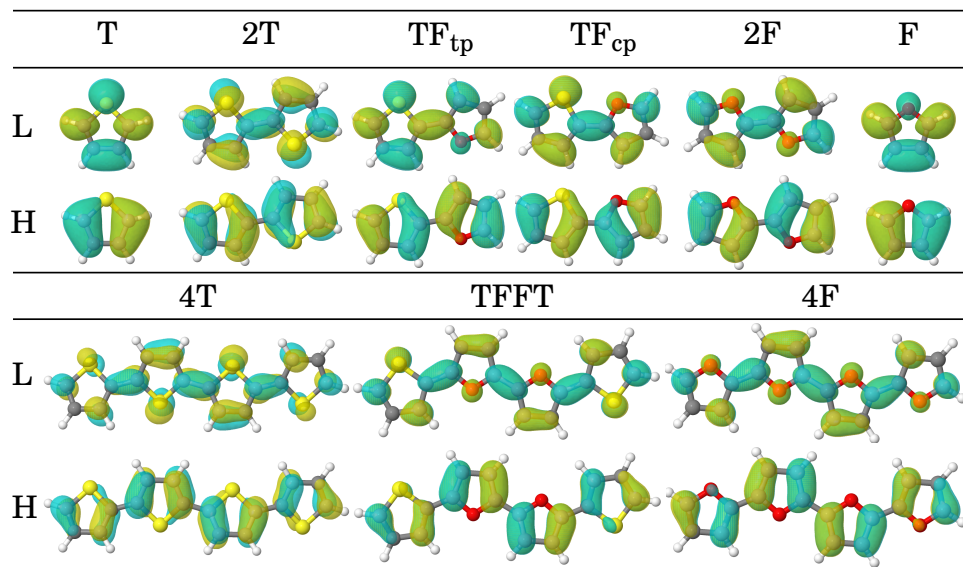


Figure 3.7: HOMO and LUMO of indicated molecules. Wavefunctions isosurface cutoff of 0.08.

In Figure 3.8 we present the optical excitations of the thiophene and furan monomers, dimers, and tetramers calculated with BSE. We indicate with blue arrows the maxima of the experimental absorption peaks reported in the literature, and for the T, F, and 2T, for which we have experimental data for the two first absorption peaks, the relative peak intensities are indicated by blue horizontal bars. The absorption spectra are for a restricted set of systems, and in Table 3.5, we detail the results for the most intense optical excitation concerning the first absorption peak for all systems investigated here. For all systems, there is only one relevant optical excitation composing the first absorption peak, which is mainly HOMO→LUMO (> 90%), except for the T monomer as already discussed. These transitions are all of $\pi \rightarrow \pi^*$ type as seen in Figure 3.7, where we present the HOMO and LUMO for a representative set of systems.

As seen in Figure 3.8, our results are in good agreement with the experimental data, with the differences concerning the optical excitations ranging from 0.01-0.15eV. Regarding photovoltaic application, the first optical excitation (first and second for the thiophene molecule) is the most relevant since it is placed in the near-ultraviolet/visible range of the solar spectrum (6.2-1.5eV). With the increasing number of chain moieties, this excitation is redshifted and placed in a better range for sunlight harvesting, achieving an energy value of around 3.0eV for the tetramers, in the violet range of visible light. As discussed before for the electronic gaps, also here, the exciton binding energies (E_B), and consequently, the optical gaps (E_{opt}) for the thiophene-furan co-oligomers are more strongly dictated by thiophene than by furan. We see this behavior by comparing the first optical excitation for the dimers in Figure 3.8, namely, the

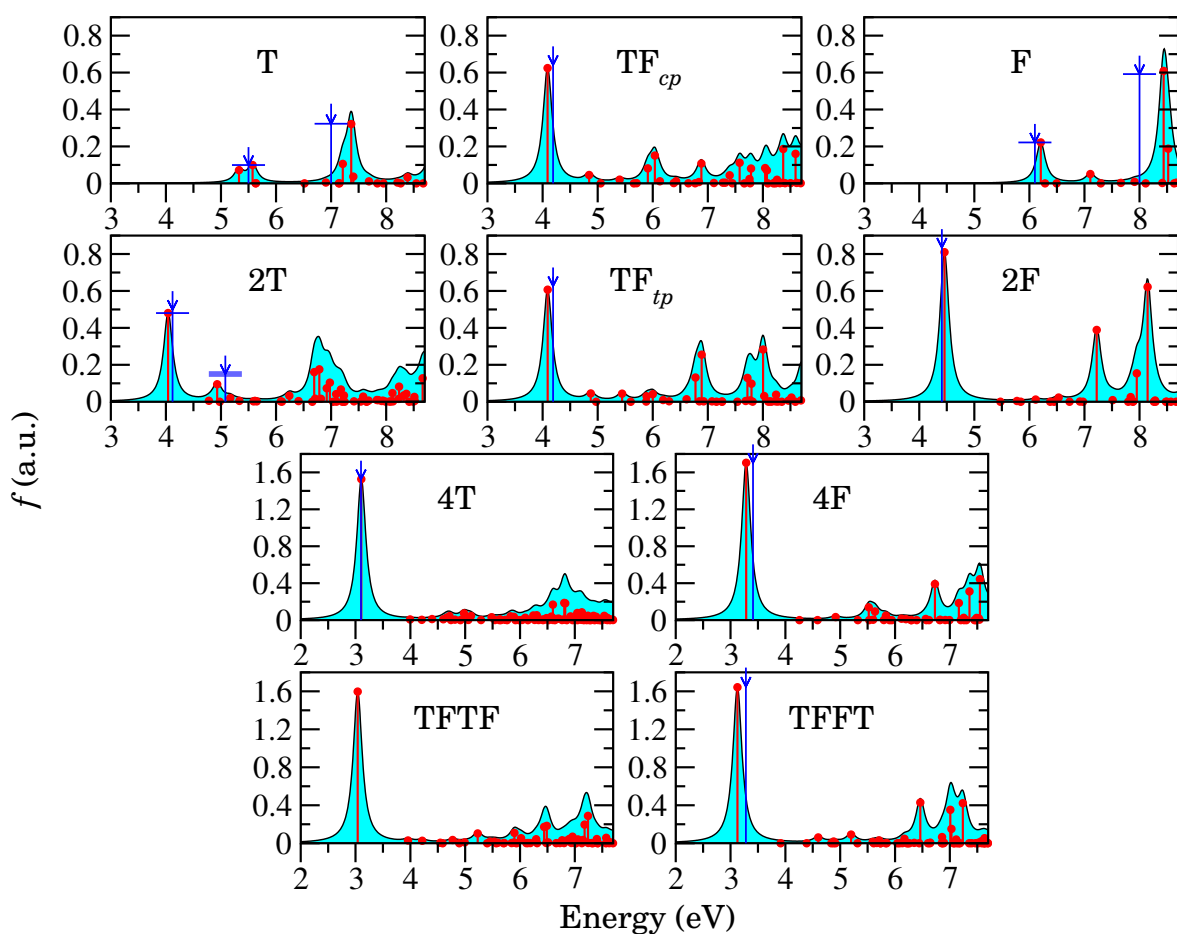


Figure 3.8: Absorption spectra for thiophene and furan monomers, dimers, and tetramers calculated with BSE/G0W0/PBE0 [92]. The red dots concern the excitation energies and the red bars are their respective oscillator strengths (f) given in arbitrary units (a.u.). The blue solid curves are an estimate of the absorption spectra of the molecules using Lorentzian functions with a scale parameter of 0.1 eV. The blue arrows indicate the position of the experimental absorption peaks (see Table 3.5) and the blue horizontal bars are an estimate of the relative peak intensities [6, 11, 81, 82, 83, 93, 94, 95].

TF_{tp} and TF_{cp} excitation energies are similar to that of 2T. Regarding the oscillator-strength values (f), we see that they increase with increasing chain size since the transition dipole moment for the HOMO \rightarrow LUMO transitions is directed along the chain direction. The f values rise sharply when moving from monomers to dimers but slow down for longer chains (f will stabilize at a certain chain size). Furthermore, concerning systems with the same chain size, the f values obtained for the thiophene and furan co-oligomers are placed between those for the thiophene (lower) and furan (higher) oligomers. These results have recently been published in a well-established journal [92].

We move now to the discussion of the cyanoacrylate end-capped oligothiophene molecules.

Table 3.5: First-relevant optical excitation for the thiophene and furan oligomers and co-oligomers calculated with BSE: exciton binding energy (E_B), optical-excitation energy (E_{opt}), and oscillator strength (f). The optical excitations are mainly HOMO \rightarrow LUMO ($\pi \rightarrow \pi^*$ type). Experimental data (E_{exp}) is presented when available.

	E_B (eV)	E_{opt} (eV)	f (a.u.)	E_{exp} (eV)
T		5.33	0.071	5.50 [81]
	4.52	5.57	0.099	
F	4.61	6.21	0.222	6.10 [93]
2T	3.68	4.04	0.481	4.09,4.12,4.29 [82, 96, 76]
TF _{tp}	3.64	4.09	0.607	4.19 [94]
TF _{cp}	3.61	4.09	0.625	4.19 [94]
2F	3.66	4.46	0.810	4.41 [95]
3T	3.18	3.41	0.989	3.50 [82]
TFT	3.16	3.41	1.062	-
FTF	3.17	3.44	1.130	-
3F	3.25	3.70	1.198	3.74 [11]
4T	2.91	3.11	1.527	3.10 [83]
TFTF	2.90	3.04	1.597	-
TFFT	2.92	3.13	1.643	3.28 [6]
4F	3.00	3.28	1.704	3.41 [11]

3.3 Cyanoacrylated oligothiophenes

Cyanoacrylate end-capped oligothiophenes (n TCs) are interesting systems since the cyanoacrylate end unit is a stronger electron-withdrawing group, modifying the optoelectronic properties of the functionalized species. n TCs are reported to exhibit a photoexcited donor-acceptor character with the HOMO localized on the oligothiophene moiety and the LUMO on the TC termination [18, 19, 20, 21]. Such donor-acceptor character provides several opportunities for applications, for example, in DSSCs [18, 19, 21, 24], since the photoexcited charge localization near the molecule-substrate interface can improve charge injection into the substrate. Furthermore, the charge displacement between the HOMO and LUMO can decrease the exciton binding energy and reduce the electron-hole recombination rate in the molecule [97]. Regarding light absorption properties, n TCs show a strong redshift in the first absorption peak in comparison with that for the non-functionalized oligothiophenes, placed in a better range for sunlight capture [19, 21, 24].

Here, we present a theoretical investigation of the structural and optoelectronic properties

of cyanoacrylate end-capped oligothiophenes with 1 and 3 thiophene units, hereafter named TC and 3TC. For 3TC, following the experimental works [18, 19, 21, 24], two kinds of structures were studied by adopting different alkylation schemes, *i.e.*, positioning and size of the alkyl groups: 1,3-OTC (first and third thiophene rings octylated) and 1,2-HTC (first and second thiophene rings hexylated).

We started by studying the geometry of the TC molecule. First, we obtained the torsion potential between the T and C moieties as shown in Figure 3.9. Similarly to the calculations in Section 3.2, geometry optimizations were performed for the TC molecule with constrained $S1'-C2'-C4''-C3''$ dihedral angle values ranging from 0 to 180° at 10° steps. The calculations were performed using PBE0/6-31g in NWChem by adopting the DRIVER module with the DEFAULT set of convergence criteria and SCF convergence energy of $1E-6$ Ha. As seen in Figure 3.9, the most stable conformation has a dihedral angle of 0° , with an energy difference of around 0.025eV ($\approx 1k_B T$) with respect to the conformation at 180° . There is a transition state at 90° with an energy of around $17.5k_B T$ about the global-minimum energy at 0° . Next, we reoptimized the geometries at 0° and 180° with 6-31g(d,p) basis set, which includes *d*-type polarization functions for carbon and sulfur atoms, to get a better description of the *S*-C bonds. These geometry optimizations were performed using PBE functional with the TIGHT convergence criteria and included the Grimme's DFT-D3 vdW correction [98]. We confirmed the energy difference of 0.025eV between the conformations at 0° and 180° and obtained bond lengths in better agreement with the results for the oligothiophenes in Section 3.1.

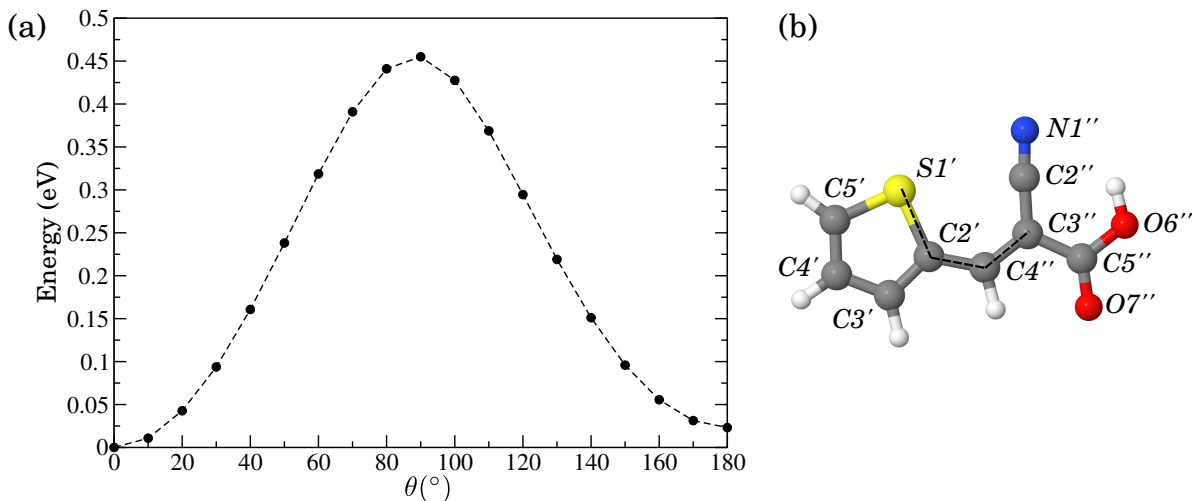


Figure 3.9: (a) TC torsion potential calculated with PBE0. (b) TC geometry: the dihedral angle θ is indicated by the dashed line, where $\theta=0^\circ$ in the corresponding figure. The global minimum energy is set to zero.

In Table 3.6 we present a comparison of the TC bond lengths obtained through geometry optimizations performed in the NWChem, FHI-aims, and QE codes, where, for these last two calculations, we adopted the NWChem final geometry as starting-point geometry. In FHI-aims

we used PBE/TIGHT-tier2 and included the Tkatchenko-Scheffler vdW (TS-vdW) correction [99]; a standard set of SCF convergence parameters was adopted, and a geometry-optimization criterion of $5\text{E-}4 \text{ eV}\cdot\text{\AA}^{-1}$ for the maximum force component. Regarding the calculation performed in QE, we used PBE/ONCVPP with a wavefunction (electronic-density) cutoff of 150Ry (600Ry) and included intracell TS-vdW correction. An SCF convergence criterion of $1\text{E-}10\text{Ry}$ for the total energy was adopted, and relaxation convergence criteria for the total energy and force components of $1\text{E-}6\text{Ry}$ and $1\text{E-}5\text{Ry}/\text{Bohr}$, respectively. For the molecule-in-a-box model, we adopted a distance of 20\AA between molecule images and also used the Martyna-Tuckerman (MT) correction [100] to truncate electrostatic interactions between the images. As we see in Table 3.6, there is a good agreement between the results obtained with the three methodologies.

Table 3.6: TC bond lengths (\AA) obtained with different methods. The bonds are identified according to Figure 3.9.

	NWChem 6-31g(d,p)	FHI-aims tier2	QE ONCVPP-150Ry
$S1' - C2'$	1.76	1.75	1.75
$C2' - C3'$	1.40	1.40	1.40
$C3' - H3'$	1.09	1.09	1.09
$C3' - C4'$	1.41	1.41	1.41
$C4' - H4'$	1.09	1.09	1.09
$C4' - C5'$	1.39	1.38	1.38
$C5' - H5'$	1.09	1.09	1.08
$C5' - S1'$	1.73	1.71	1.71
$C2' - C4''$	1.43	1.42	1.42
$N1'' - C2''$	1.18	1.17	1.17
$C2'' - C3''$	1.42	1.42	1.42
$C3'' - C4''$	1.38	1.37	1.37
$C4'' - H4''$	1.10	1.09	1.09
$C3'' - C5''$	1.51	1.51	1.51
$C5'' - O6''$	1.36	1.36	1.36
$O6'' - H6''$	0.98	0.98	0.97
$C5'' - O7''$	1.22	1.21	1.21

Moving to the geometry of the 3TC systems, we first discuss the structure of the 1,3-OTC. From the results presented in Section 3.1, the 3T more stable conformer exhibits a *trans*-torsioned conformation with *S-C-C-S* dihedral angle around 150° . In this respect, as presented in Figure 3.10-a and b, we considered two possible conformations for the 1,3-OT moiety, based on the alkylation scheme proposed in Ref. [21]. In Figure 3.10-a the thiophene side rings are functionalized with octyl side chains and are alternated, as more clearly seen in the side view.

In this case, the 1,3-OT moiety displays a C_2 -symmetry axis crossing the central thiophene ring. As another possibility, in Figure 3.10-b, the thiophene side rings functionalized with octyl side chains are parallel. The 1,3-OT now displays a σ_s -symmetry plane crossing the central thiophene ring. Both structures were relaxed in NWChem using PBE/6-31g(d,p) with the TIGHT convergence criteria and constrained point-group symmetry. We found that the C_2 -symmetry structure is 0.047eV ($\approx 1.8k_B T$) more stable than the C_s -symmetry structure, where the difference concerns the alkyl side chains. Regarding the $S-C-C-S$ dihedral angle, we obtained a value of 148.4° for the C_2 -symmetry structure and 146.2° for the C_s -symmetry structure. Both are smaller than the value obtained for the non-alkylated 3T system in Section 3.1 with the PBE0 functional (around 160°).

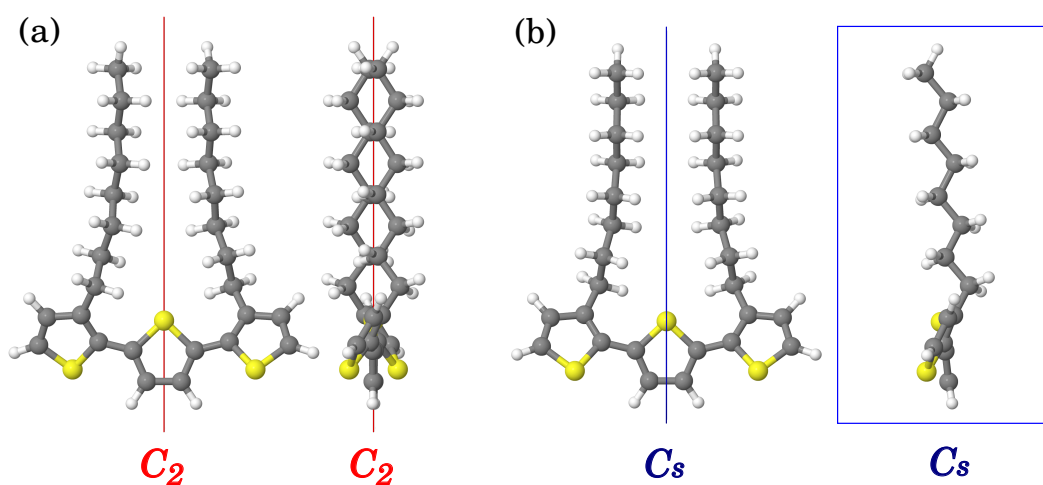


Figure 3.10: Top and side views of the C_2 -symmetry (a) and C_s -symmetry (b) structures of 1,3-OT obtained through PBE/6-31g(d,p) symmetry-constrained relaxation with the DFT-D3 vdW correction in NWChem.

The 1,3-OTC starting-point geometry was constructed based on the more stable conformer of TC, *i.e.*, with $S1'-C2'-C4''-C3''$ dihedral angle value of 0° , and the C_2 -symmetry structure of 1,3-OT. The geometry was optimized using PBE/TIGHT-tier2 and the TS-vdW correction in FHI-aims, adopting the same optimization criteria used for the TC. In Figure 3.11-a we present the final geometry of 1,3-OTC. As we see, the general structure of the 1,3-OT and TC moieties is maintained. We further explored another proposition [19] in which the carbon atoms of the octyl side chains are all lying in the plane of the molecule. In this case, presented in Figure 3.11-b, by starting with a geometry with $S-C-C-S$ dihedral angles close to 180° , we obtained a planar structure, with energy 0.24eV ($\approx 9.3k_B T$) higher than that for the non-planar structure in Figure 3.11-a.

Another proposition [18, 19, 24], named here 1,2-HTC, has a different alkylation scheme as presented in Figure 3.12, where the first and second thiophene rings are functionalized with hexyl side chains and the cyanoacrylated ring has no hexyl functionalization. In Figure 3.12-a

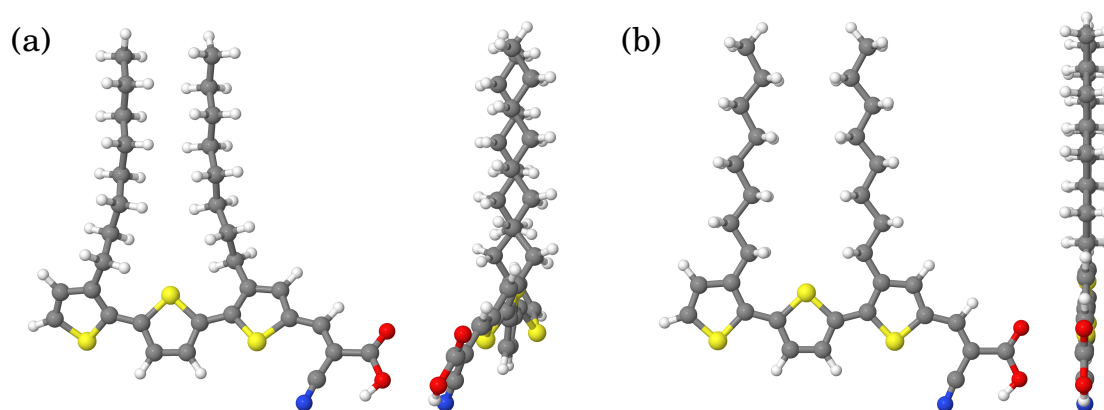


Figure 3.11: Top and side views of the non-planar (a) and planar (b) structures of 1,3-OTC obtained through PBE/TIGHT-tier2 relaxations with the TS-vdW correction in FHI-aims.

we present the geometry of the 1,2-HT moiety and in b the geometry of the 1,2-HTC. Both systems were relaxed using PBE/TIGHT-tier2 in FHI-aims, adopting the same optimization criteria as before for the SCF cycles, and a geometry-optimization criterion for the maximum force component of $1\text{E-}3 \text{ eV}\cdot\text{\AA}^{-1}$. The 1,2-HT geometry displays, from left to right, *S-C-C-S* dihedral angle values of 149.5° and 148.2° . After functionalization with the cyanoacrylate group, these values increase to 160.1° and 170.7° . By further including the TS-vdW correction in the relaxation of the 1,2-HTC, no significant changes are seen in these dihedral angle values, only a small modification in the *S-C-C-C* dihedral angle between 1,2-HT and C, moving from 0.2° to 1.4° . In fact, the vdW is more important for the conformation of the 1,3-OTC due to the interaction between the octyl side chains which are closer in that case, since the octylated thiophene mers point in the same direction.

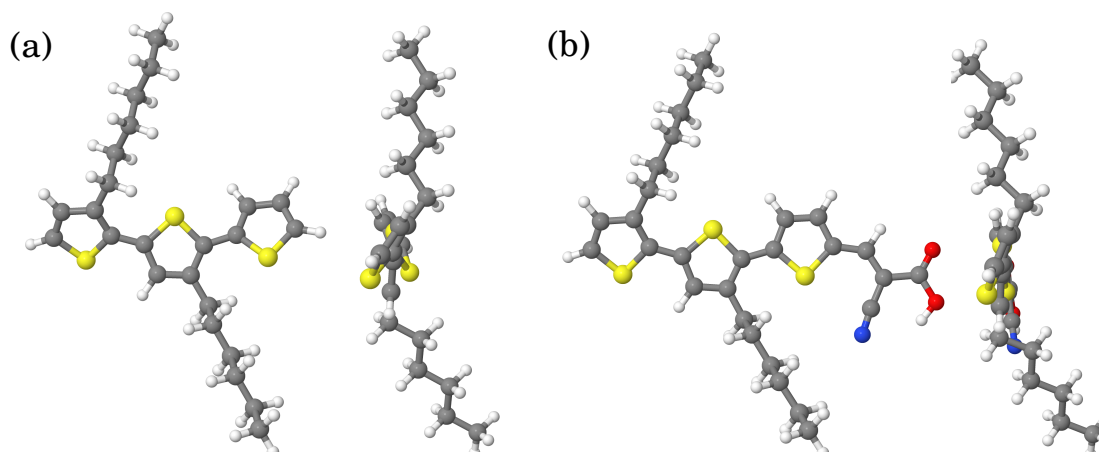


Figure 3.12: Top and side views of the geometries of 1,2-HT (a) and 1,2-HTC (b) obtained through PBE/TIGHT-tier2 relaxations in FHI-aims.

At this point, we move on to the electronic structure properties of the TC and 3TC sys-

tems. Initially, we observed that the alkyl side chains in 1,3-OTC and 1,2-HTC have just a small contribution to the frontier electronic states of the HOMO-LUMO gap. Since we are interested in this energy range, the following results were obtained for ethylated systems, *i.e.*, 1,3-ETC and 1,2-ETC, by shortening the alkyl side chains and keeping the final geometries of 1,3-OTC and 1,2-HTC. The electronic structure of TC and 1,2-ETC systems was also studied in QE. In this case, the 1,2-ETC geometry was reoptimized using PBE/ONCVPP in QE, adopting the same settings described above for the TC and a 20Å distance between molecule images. The electronic structure results were obtained at the Γ -point using PBE0/ONCPP with wavefunction (electronic density) cutoff energy of 90Ry (360Ry) and, concerning the Fock-operator expansion, cutoff energy of 360Ry (180Ry) for the TC (1,2-ETC). We further included the MT correction and adopted an SCF convergence criterion of 1E-10Ry and 1E-6Ry concerning the total energy for the TC and 1,2-HTC, respectively.

In Figure 3.13 we show the energy levels for TC and 1,2-ETC around the HOMO-LUMO gap region (from HOMO-10 to LUMO+10), calculated with FHI-aims and QE. There is a good agreement between both methodologies from the HOMO-10 to LUMO+2 energy levels for both systems. The orbital electronic charge densities also have a similar character across all these states, starting to deviate at the LUMO+3 since unoccupied states are not self-consistently converged. This correspondence supports our results and is particularly relevant for the discussion in Section 4.2, where we present an investigation performed with the QE of these molecules attached to a substrate.

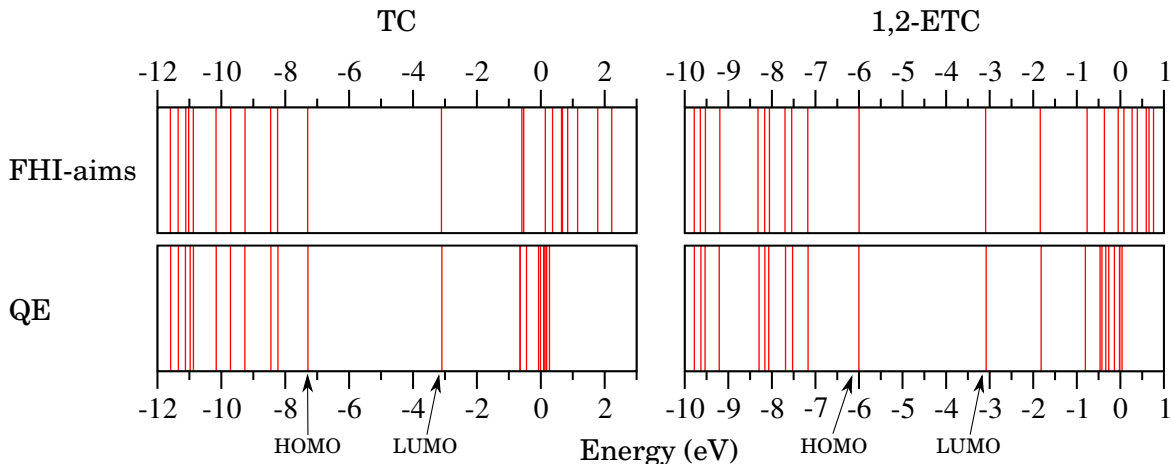


Figure 3.13: HOMO-LUMO gap frontier energy levels for TC and 1,2-ETC calculated using PBE0/NAO-VCC-3Z in FHI-aims (top) and PBE0/ONCVPP in QE (bottom): energy levels ranging from HOMO-10 to LUMO+10. The energies are relative to the vacuum level.

In Table 3.7 we present the IP and EA values for TC, 1,3-ETC, and 1,2-ETC obtained with bare PBE0 and including the GW correction in FHI-aims. We observe that the IP_{GW} values are similar to those obtained for the oligothiophenes presented in Table 3.4, concerning systems with the same number of thiophene rings. This result comes as a consequence of the influence

of thiophene in the HOMO state of the n TCs, as will be discussed below. Another interesting aspect is the similarity between the EA values of the TC and 3TC systems obtained with bare PBE0, which indicates a similar character to the LUMO states.

Table 3.7: GW ionization potential (IP_{GW}), electron affinity (EA_{GW}), and electronic gap (E_{gGW}) of TC and 3TC calculated in FHI-aims. The Kohn–Sham IP values (IP_{KS}) and EA values (EA_{KS}) obtained with PBE0 are also presented.

	IP_{KS}	IP_{GW}	EA_{KS}	EA_{GW}	E_{gGW}
TC	7.30	8.74	3.11	1.66	7.08
1,3-ETC	6.10	7.18	3.10	1.97	5.21
1,2-ETC	6.00	7.05	3.09	2.00	5.05

In Figure 3.14 we show the GW energy level diagram for 1,2-ET, 1,2-ETC, and TC. The unfolding of the 1,2-ET and TC orbitals into the orbitals of the 1,2-ETC is indicated by dotted lines. As commented before, the 1,2-ETC HOMO state is dictated by thiophene. On the other hand, the 1,2-ETC LUMO is dictated by the TC moiety. The same behavior is observed for the 1,3-ETC which presents a similar electronic structure to 1,2-ETC, which justifies the similarity between the EA values presented in Table 3.7. Indeed, the LUMO state of the n TCs is strongly influenced by the cyanoacrylate group, which acts as an electron attractor center, bringing down the energy of the LUMO state. For example, the TC presents an IP_{GW} similar to T, namely 8.74eV and 8.71eV, respectively, however, a significantly smaller E_{gGW} of 7.08eV in comparison with 10.09eV for the T.

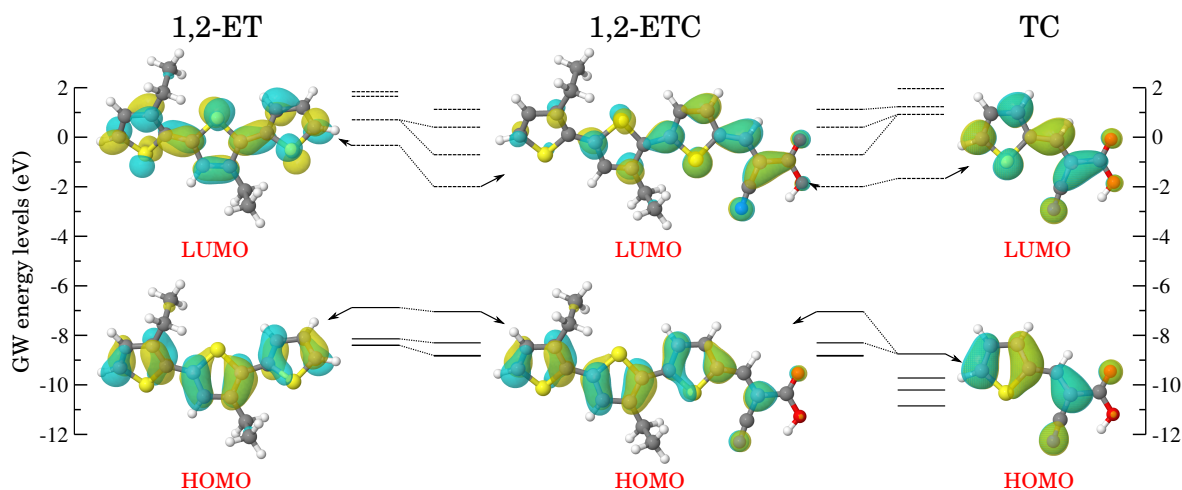


Figure 3.14: GW energy level diagram for 1,2-ET, 1,2-ETC, and TC. The dotted lines indicate the unfolding of the 1,2-ET and TC orbitals into the orbitals of the 1,2-ETC. Wavefunctions isosurface cutoff of 0.08.

Regarding the optical properties, in Figure 3.15 we present the optical excitations of TC, 1,3-ETC, and 1,2-ETC calculated with BSE in FHI-aims. The calculated optical excitation

energies and respective oscillator strengths are in red, while the blue arrows indicate the experimental data for the maximum absorption peak position. We see a good agreement between the theoretical and experimental results, as more clearly seen in Table 3.8, where we present the results concerning the first-bright optical excitation for each system and the experimental reference values. We observe that the functionalization with the cyanoacrylate group leads to a strong redshift in this first excitation in comparison with the results obtained for the bare oligothiophenes presented in Table 3.5. This optical excitation is mainly HOMO→LUMO for all systems, *i.e.*, TC, 1,3-ETC, and 1,2-ETC. As discussed before, the cyanoacrylate end group brings down the energy of the LUMO state, leading to the optical gap closure seen for the *n*TCs.

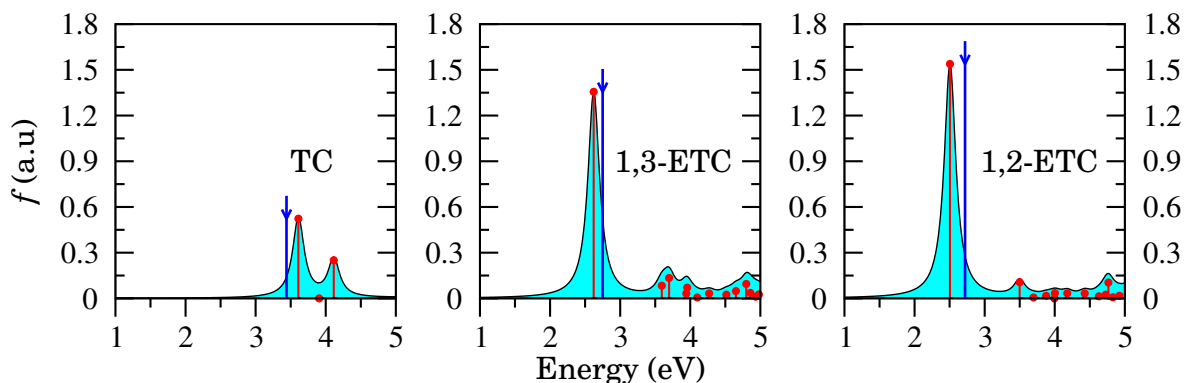


Figure 3.15: Absorption spectra for TC, 1,3-ETC, and 1,2-ETC calculated with BSE in FHI-aims. The red dots concern the excitation energies and the red bars are their respective oscillator strengths (f) given in arbitrary units (a.u.). The blue solid curves are an estimate of the absorption spectra of the molecules using Lorentzian functions with a scale parameter of 0.1 eV. The blue arrows indicate the position of the experimental absorption peaks [19, 21] (see Table 3.8).

Table 3.8: First-bright optical excitation for the *n*TCs calculated with BSE in FHI-aims: exciton binding energy (E_B), optical-excitation energy (E_{opt}), and oscillator strength (f). The optical excitations are mainly HOMO→LUMO ($\pi \rightarrow \pi^*$ type). Experimental data (E_{exp}) is presented.

	E_B (eV)	E_{opt} (eV)	f (a.u.)	E_{Texp} (eV)
TC	3.47	3.61	0.523	3.44 [19]
1,3-ETC	2.59	2.62	1.356	2.75 [21]
1,2-ETC	2.55	2.50	1.538	2.72 [19]

Interestingly, since the HOMO and LUMO are localized on different sides in the *n*TCs, the exciton binding energies (Table 3.8) are smaller than those obtained for the bare oligothiophenes, which will favor the exciton dissociation in these systems [97]. This charge displacement between the HOMO and LUMO also increases the transition dipole moment along the chain direction, leading to a significant increase in the oscillator strength values (f) in comparison with the bare oligothiophenes. Remarkably, in the case of the T and TC molecules, the

f value scale is more than five times, moving from 0.099 to 0.523. Furthermore, the redshift seen in the first optical excitation increases the portion of the solar spectrum swept by the n TCs. The maximum absorption peak of TC is placed in the near ultraviolet range, and for the 3TC systems, the absorption spectrum is already placed in the visible range of the solar spectrum.

The n TCs thus present promising optoelectronic properties and especially incorporate essential characteristics for the application as sensitizers in DSSCs. These properties motivated us to investigate these systems as sensitizers for the wurtzite ZnO crystal [24]. In the next chapter, we present a theoretical investigation of this setup.

Thiophene as dye sensitizer for solar cells

As discussed in the previous section, n TCs exhibit favorable properties for application as sensitizers in DSSCs. Recently, Oehrlein *et al.* [24] have shown fluorescence quenching for n TCs ($n=3, 4,$ and 5) upon binding to ZnO nanocrystals, indicating molecule-crystal photoexcited charge transfer with possible application in DSSCs. Although n TCs functionalization has already been proposed in TiO_2 -based solar cells [18, 19, 23], in the last few years we have seen a growing interest in the use of ZnO as the semiconducting layer for devices [27, 50, 101], including photovoltaic devices [22, 28]. There are several reasons for this, coming from properties as high electron mobility, direct and wide band gap of around 3.4eV [27], high-quality crystal growth by a variety of methods [27, 28], high natural abundance of zinc [50], biocompatibility (it does not cause an adverse effect in living organisms [101]), application in water-soluble devices [101], and so forth.

In this chapter, we present a theoretical investigation of the ZnO ($10\bar{1}0$) surface functionalized with the 1,2-ETC molecule, as suggested in Ref. [24]. We start with the structural and electronic structure properties of the ZnO crystal and the chosen surface, and then move to the discussion of the anchored system.

4.1 ZnO crystal and ($10\bar{1}0$) surface

In Figure 4.1 we show the wurtzite ZnO crystal unit cell adopted here. It presents an hexagonal structure where $|a_1| = |a_2| = a$, with a_1 directed along the x axis, a_2 forming an angle of $2\pi/3$ with a_1 in the xy plane, and $a_3 = c$ along the z axis. The atomic basis has two ZnO dimers with the respective $Zn-O$ bonds along a_3 , as shown in Figure 4.1-b, where one Zn atom is placed at the bottom of the unit cell and the other at half-cell. Looking from the top along the z direction, as shown in Figure 4.1-a, each ZnO dimer axis is placed on the centroid of an equilateral triangle of side a , formed by straight lines joining the vertices of the hexagonal unit cell.

Here, we perform a DFT study of the ZnO crystal and ($10\bar{1}0$) surface with the planewave-based method QE (see Section 2.5.2). The crystal was fully relaxed using PBE/ONCVPP and an energy cutoff for the wavefunctions (electronic density) of 150Ry (600Ry). The Hubbard-U correction ($U=8.5eV$) was included for the Zn $3d$ -states. Regarding the \vec{k} -vectors in the BZ,

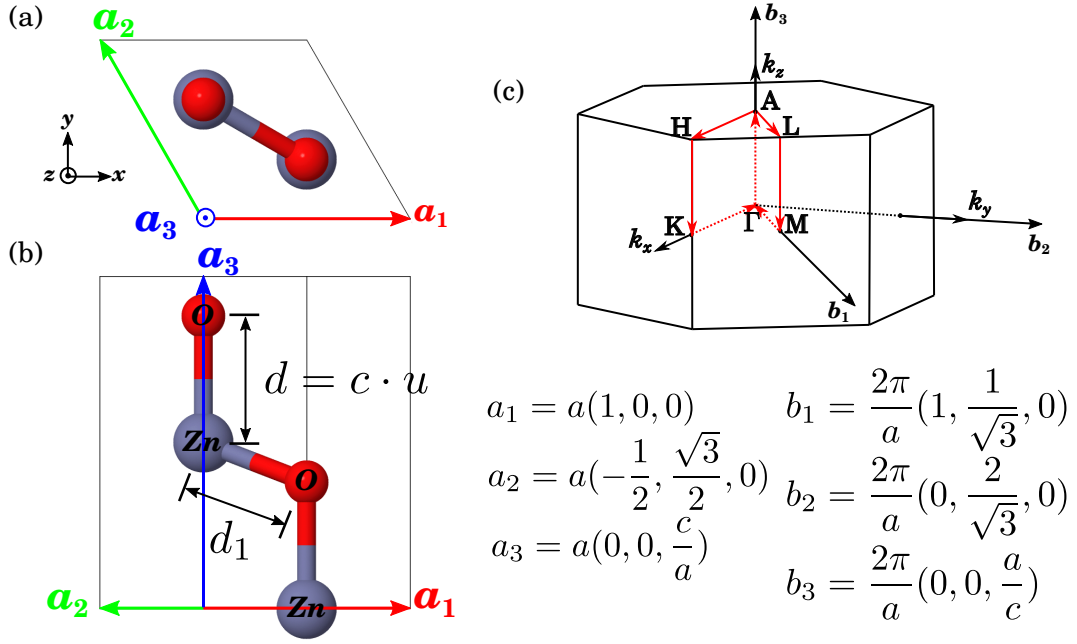


Figure 4.1: Top (a) and side (b) views of the *ZnO* crystal unit cell, and respective BZ (c), adapted from Ref. [102], including the path used for the band structure visualization.

we adopted a Monkhorst-Pack grid of $10 \times 10 \times 10$ centered at the Γ -point. This setup ensured a total-energy convergence of $1\text{E-}4\text{Ry}$. For the SCF cycles, we adopted a total-energy convergence threshold of $1\text{E-}10\text{Ry}$, and in geometry optimizations, total-energy and maximum-force-component convergence thresholds of $1\text{E-}6\text{Ry}$ and $1\text{E-}5\text{Ry/Bohr}$, respectively.

In Table 4.1 we present the results obtained for the structural parameters indicated in Figure 4.1 and reference values from a HF/LCAO calculation and X-Ray Diffraction (XRD) measurement [103]. As would be expected, the HF method gives a *Zn-O* bond length along \mathbf{a}_3 (d value) smaller than the PBE result, which leads to the difference between the c values. Such a trend is not verified for the *Zn-O* bond length in the diagonal direction (d_1 value), which is the same with both methods. By comparing our PBE result with the measured values, we see a deviation for a and c , however, this corresponds to a slight difference for the *Zn-O* bond lengths d and d_1 of 0.03\AA and 0.02\AA , respectively. Accordingly, we can consider our results to be in good agreement with the reference values. It is worth noting the very small difference between d and d_1 in the theoretical results of only 0.01\AA , exhibiting practically the same length, as indicated in the experimental data.

In Figure 4.2 we present the band structure and Projected Density of States (PDOS) for the *ZnO* crystal calculated using PBE/ONCVPP. For the PDOS, the Hamiltonian of the system was re-diagonalized using a denser k -points grid of $14 \times 14 \times 14$ through a non-SCF calculation. The PBE results predict a direct electronic gap at the Γ -point of 1.19eV , much smaller than the most accepted value of around 3.4eV [27, 78, 104]. At the bottom of the CB, the major contribution comes from the *Zn 4s*-states. These states exhibit a parabolic band structure around the Γ -point

Table 4.1: Structural parameters of the ZnO crystal (Figure 4.1) obtained through geometry optimization with PBE/ONCVPP in QE and reference values: “ a ” for the lattice vectors a_1 and a_2 , “ c ” for a_3 ; “ d ” (“ u ”) is the $Zn-O$ bond length along a_3 (as a ratio of “ c ”), and “ d_1 ” is the $Zn-O$ bond length in the diagonal direction.

Parameters	PBE/ONCVPP	HF/LCAO *	Exp. XRD*
a (Å)	3.29	3.29	3.25
c (Å)	5.29	5.24	5.20
c/a	1.61	1.59	1.60
u	0.38	0.38	0.38
d (Å)	2.01	1.99	1.98
d_1 (Å)	2.00	2.00	1.98

* Taken from Ref. [103].

which is characteristic of a system with high electron mobility. Since the bands are relatively sparse in this energy region, we see a low density of states in the PDOS, with a peaked character around the inflection points. As commented in Section 2.4, the Hubbard-potential correction brings on the separation of the O $2p$ (in blue) and Zn $3d$ (in green) bands seen in the PDOS. Considering the maximum point in the Zn $3d$ -band, it is placed at -8.1eV from the VBT, in good agreement with the reference values [50, 105].

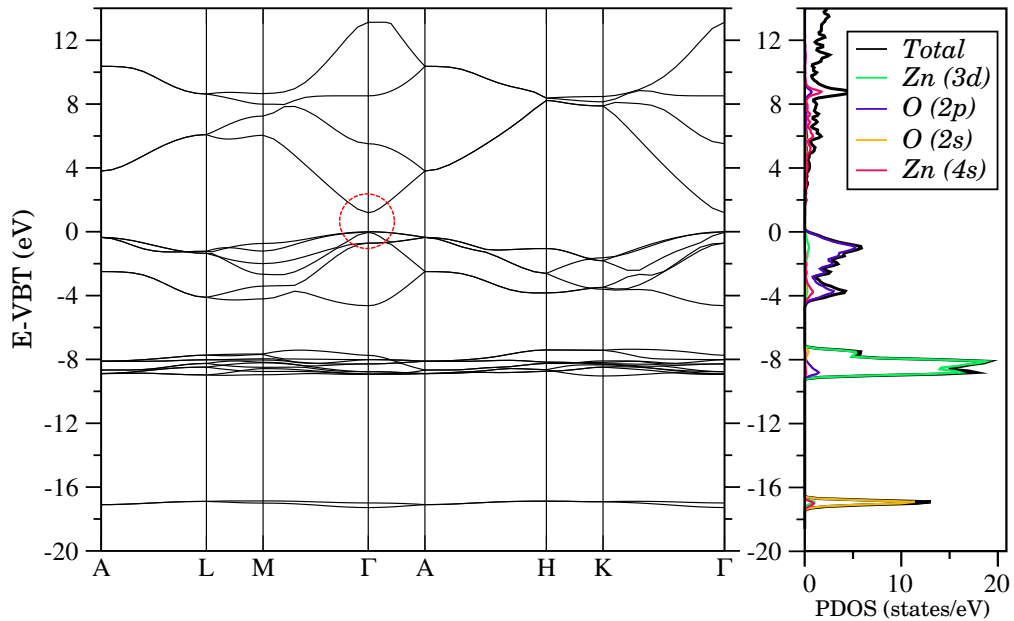


Figure 4.2: Band structure and PDOS for the ZnO crystal obtained with PBE/ONCV in QE. A Gaussian broadening of 0.01Ry was adopted for the PDOS. The red circle indicates the position of the gap in the BZ. VBT set to zero.

As well known, hybrid-GGA functionals are more reliable than pure-GGA to study the

electronic structure properties of semiconductors [27, 50]. Therefore, we further performed a PBE0 calculation to investigate the frontier electronic states around the ZnO gap. In Figure 4.3 we show the band structure of the ZnO calculated with PBE and PBE0 using a k -points grid of $16 \times 16 \times 16$, and concerning the Fock-operator expansion, cutoff energy of 300Ry. Indeed, the inclusion of exact exchange leads to gap opening, with a final value of 3.56eV in our results.

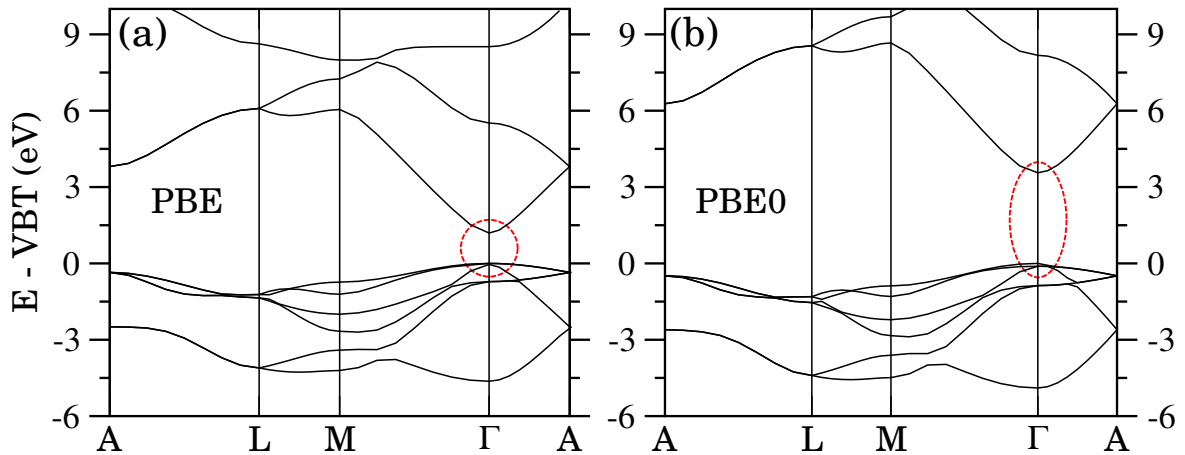


Figure 4.3: Band structure for the ZnO crystal calculated with PBE (a) and PBE0 (b). The red circle/ellipse indicates the position of the gap in the BZ. VBT set to zero.

We investigate the ZnO ($10\bar{1}0$) surface that, as reported in the literature, is the most stable surface of the wurtzite ZnO crystal and the dominant surface in many ZnO nanostructures [25, 26]. In this case, we adopted the slab model to simulate a surface [25, 106, 107], and the unit cell is shown in Figure 4.4. We used an orthorhombic lattice with a slab containing 6 ZnO layers, as indicated in Figure 4.4-b, and a vacuum layer of 35\AA between the slabs, along the z direction, to ensure isolation.

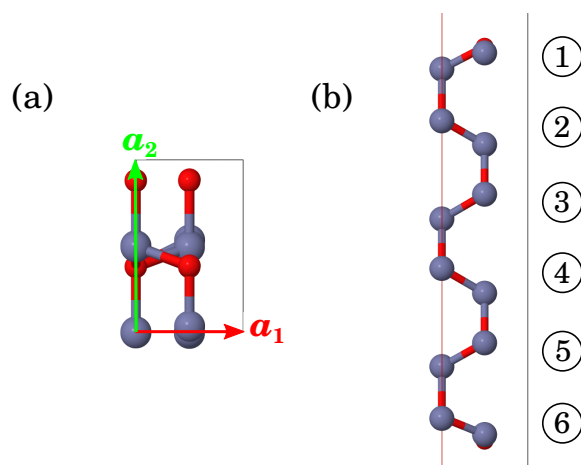


Figure 4.4: Top (a) and side (b) view of the ZnO ($10\bar{1}0$) surface unit cell, with labeled layers in (b). $a_1=3.3\text{\AA}$ and $a_2=5.3\text{\AA}$. A vacuum layer of 35\AA was adopted between the slabs along the z direction.

We started with a slab containing 8 *ZnO* layers and with double lattice constant along a_1 , built from the *ZnO* crystal final geometry and adopting a vacuum layer of 30Å between the slabs along the z direction. This starting-point structure was relaxed using PBE/ONCVPP at the Γ -point only, adopting the same optimization settings as for the crystal. During geometry optimization, we kept fixed the supercell lattice vectors and the four central *ZnO* layers to simulate the bulk effect. Starting from this optimized structure, we then built the unit cell shown in Figure 4.4 by removing the two central *ZnO* layers and using a single cell along a_1 . This allowed us to adopt a denser k -points grid of 10x10x1 in the subsequent SCF calculation with PBE, for which the results are presented below.

In Figure 4.5 we show the PDOS for the *ZnO* crystal and (10 $\bar{1}$ 0) surface (aligned through the *O* 2*s*-bands) which exhibit a similar character. Even the band gap is almost the same despite the broken symmetry, namely 1.2eV for the crystal and 1.19eV for the surface. We see a band enlargement for the surface which is more pronounced in the *O* 2*s*-band. While the DOS on the two central layers (3 and 4) behaves similarly to the bulk, we see a growing dispersion around this result as we move to the final layers: the DOS on the 2 and 5 layers are the same and show a low dispersion around the bulk signal, and the DOS on the 1 and 6 layers are also identical with a higher dispersion around the bulk signal. The band structure of the (10 $\bar{1}$ 0) surface is shown in Figure 4.6. The number (and relative energy displacement) of parabolic bands in the region of the bottom of the conduction band comes from the electronic confinement in the finite slab model, and as such, even naming it as CBM, CBM+1, and CBM+2, these are states at the low energy region of the band, which continues to show good electron mobility character.

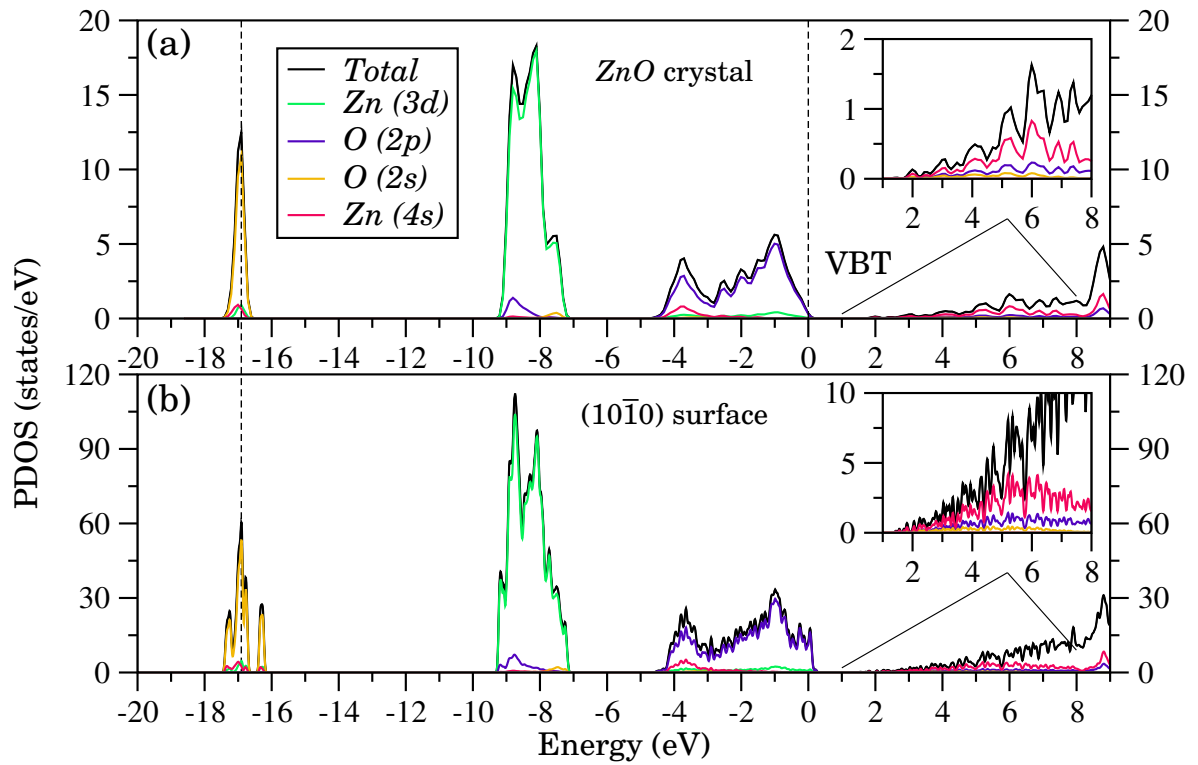


Figure 4.5: PDOS for the ZnO crystal (Gaussian broadening of $0.01Ry$) (a) and $(10\bar{1}0)$ surface (Gaussian broadening of $0.003Ry$) (b), calculated with PBE/ONCVPP. The dashed lines indicate the crystal VBT, set to zero, and the alignment of the crystal and surface $O\ 2s$ -band. In the inset a detailed view of the CB close to the gap region.

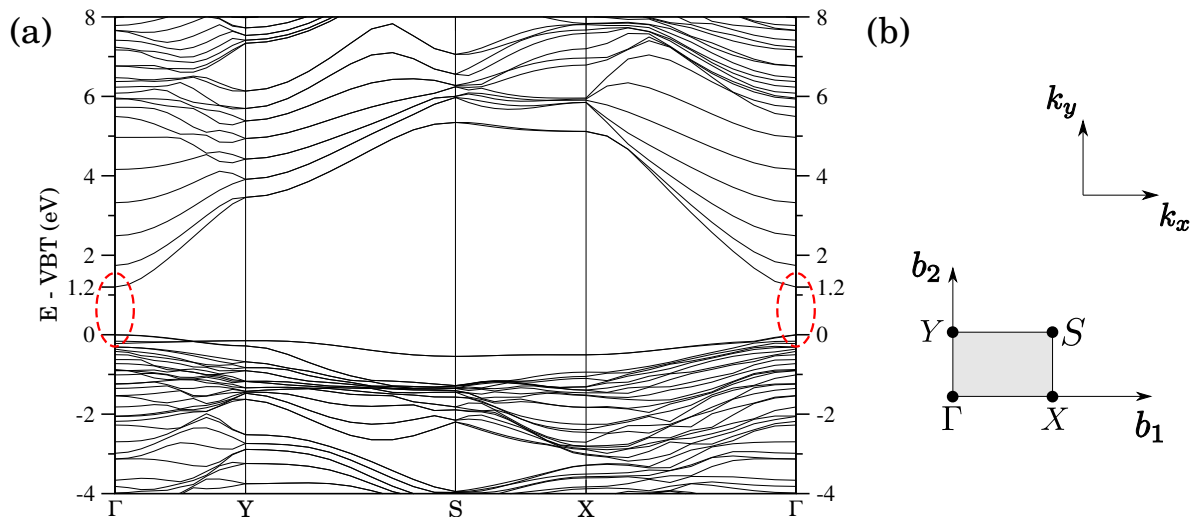


Figure 4.6: (a) Band structure for the $(10\bar{1}0)$ surface calculated with PBE/ONCVPP. The dashed ellipses indicate the position of the electronic gap in the BZ. VBT set to zero. (b) BZ picture with the high-symmetry points adopted in the band structure path.

4.2 ZnO ($10\bar{1}0$) surface functionalized with nTC s

In this section, we present a theoretical investigation of the electronic structure properties of the ZnO ($10\bar{1}0$) surface sensitized with nTC molecules ($nTC-ZnO$). The geometry optimizations and electronic structure preliminary calculations were performed by adopting DFT with PBE/ONCVPP in QE. Next, we investigated the electronic structure properties using PBE0/ONCVPP. In the calculations, again a Hubbard-U correction of 8.5eV was included for the Zn 3d-states, and intra-cell TS-vdW correction for the geometry optimizations, *i.e.*, accounting only for the charge interactions inside the same cell.

A surface supercell with dimensions of $4 \times 2 \times 1$ was adopted to simulate a single isolated molecule anchored to the surface, as shown in Figure 4.7. We start by investigating the attachment of the TC molecule ($TC-ZnO$) since it is the terminal group in longer nTC s. The molecule was attached to a surface Zn site following the mechanism suggested by Dominguez *et al.* [25]. The oxygen atom of the $-OH$ radical is predicted to form a single bond with a surface Zn site through hydrogen-atom dissociation by a first-neighbor oxygen site at the surface. We arranged the molecule in such a way that the $=O$ radical could also form a bond with another surface Zn site during geometry optimization. A vacuum layer of 25\AA was adopted along the z direction to simulate an isolated system.

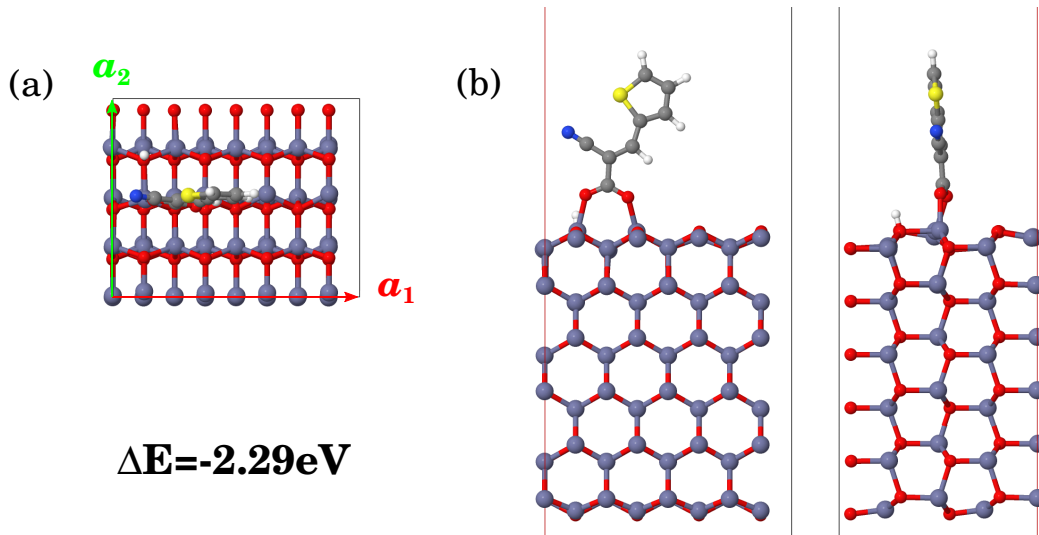


Figure 4.7: Top (a) and side (b) views of the $TC-ZnO$ system. $a_1=13.2\text{\AA}$ and $a_2=10.6\text{\AA}$. A vacuum layer of 25\AA was adopted along the z direction. The molecule-surface ΔE is shown, indicating a covalent bond.

The $TC-ZnO$ geometry was optimized at the Γ -point using PBE and an energy cutoff for the wavefunctions (electronic density) of 150Ry (600Ry). A total-energy and maximum-force-component convergence thresholds of $1E-4Ry$ and $1E-3Ry/Bohr$ were adopted for geometry optimization, and a total-energy convergence threshold of $1E-8Ry$ in the SCF cycles. During

optimization, we fixed the lattice vectors and the four bottom atomic layers of the slab. We obtained a molecule-surface binding energy (ΔE) of -2.29eV , as shown in Figure 4.7, indicating chemical bonding. This value is comparable with DFT binding energies obtained for other chemically-bonded systems presented in the literature [23, 108, 109].

We used the TC-ZnO final geometry as the starting point to build the 1,2-ETC-ZnO geometry, now with a cell dimension along the z direction which ensured a vacuum layer of 20\AA for the 1,2-ETC-ZnO system. The geometry was then relaxed with the same optimization settings adopted before for the TC-ZnO system. In Figure 4.8 we show the 1,2-ETC-ZnO final geometry and highlight the molecule-surface binding energy of -2.26eV . Arbouch *et al.* [23] presented a DFT study of carboxylic-terminated short oligothiophenes anchored to an anatase TiO_2 model cluster, exposing the TiO_2 (101) surface. They show that ΔE becomes more negative with decreasing size of the thiophene conjugated part, approaching the binding energy of the sole anchor group, which is in accordance with our results.

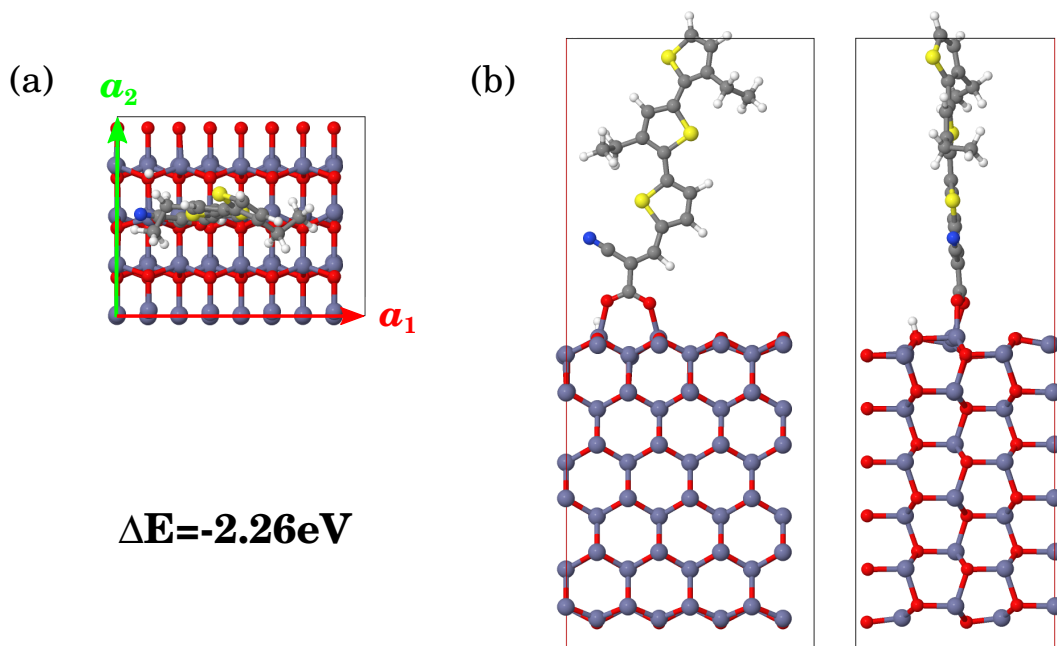


Figure 4.8: Top (a) and side (b) views of the 1,2-ETC-ZnO system. $a_1=13.2\text{\AA}$ and $a_2=10.6\text{\AA}$. A vacuum layer of 20\AA was adopted along the z direction. The molecule-surface ΔE is shown, indicating a covalent bond.

In Figure 4.9 we show the PDOS for the isolated 1,2-ETC molecule, the 1,2-ETC-ZnO attached system, and the ZnO ($10\bar{1}0$) surface, obtained with PBE. Regarding the 1,2-ETC-ZnO electronic structure calculation, we adopted a k -points grid of $3\times 5\times 1$, an energy cutoff for the wavefunctions (electronic density) of 100Ry (400Ry), and a total-energy convergence threshold of $1\text{E-}6\text{Ry}$ for the SCF cycles. For the PDOS, the Hamiltonian of the 1,2-ETC-ZnO system was re-diagonalized using a denser k -points grid of $4\times 7\times 1$ through a non-SCF calculation. The ZnO ($10\bar{1}0$) surface VBT is set to zero in Figure 4.9-c and the 1,2-ETC-ZnO and the surface PDOS in

Figure 4.9-b are aligned through the Zn 3s-bands (the contribution from the two-central atomic layers in the ZnO slab). The PDOS for the 1,2-ETC (Figure 4.9-a) was aligned to that of 1,2-ETC- ZnO through deep valence states mainly due to carbon and sulfur atoms.

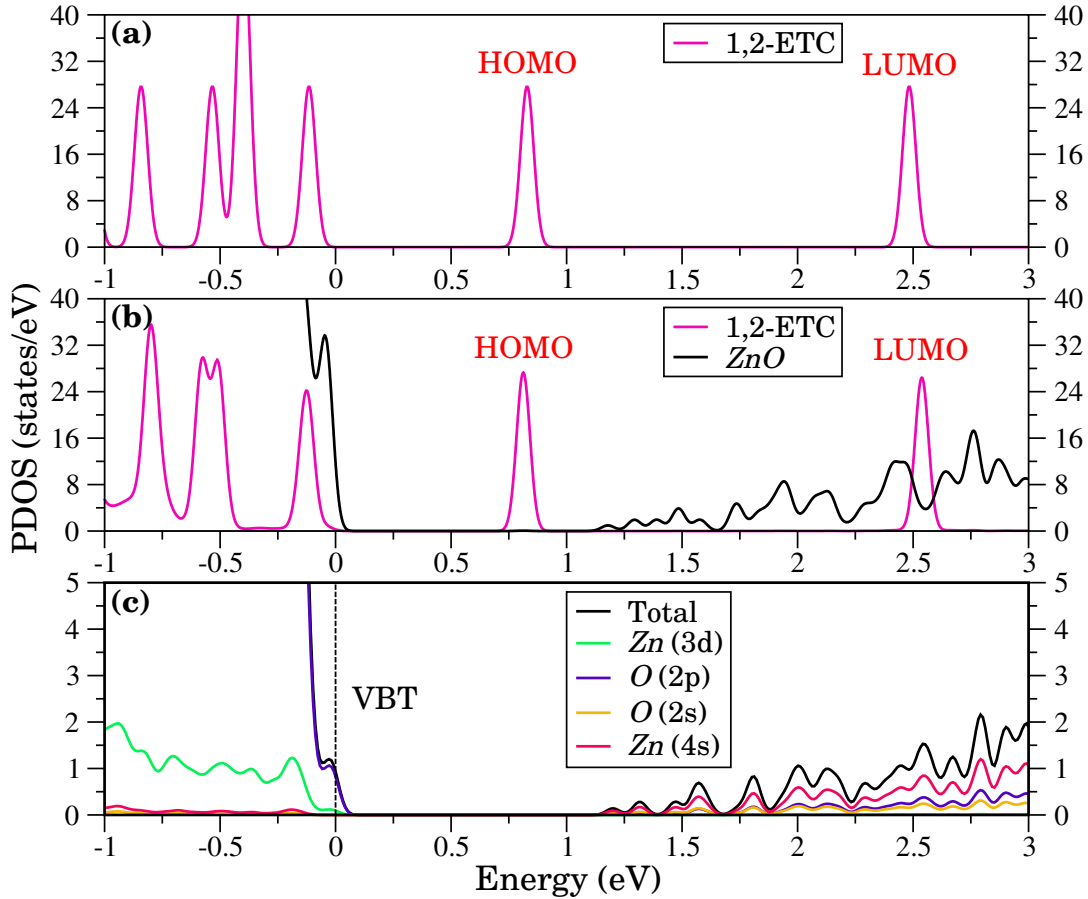


Figure 4.9: PDOS for (a) the isolated 1,2-ETC molecule, (b) 1,2-ETC- ZnO attached system, and (c) ZnO ($10\bar{1}0$) surface, calculated using PBE/ONCVPP. A Gaussian broadening of 0.003Ry was adopted in the PDOS. The DOS “a” and “b” were aligned through deep valence states mainly due to sulfur and carbon atoms, and “b” and “c” through the Zn 3s-band (contribution from the two central ZnO layers). The dashed line indicates the crystal VBT set to zero.

Focusing on the 1,2-ETC- ZnO system, as seen in Figure 4.9-b, the 1,2-ETC-HOMO energy is placed in the oxide gap, close to the CBM, whereas the 1,2-ETC-LUMO energy is 1eV above the CBM. In Figure 4.10 we show the electronic charge densities at the Γ -point concerning these levels and compare them with the HOMO and LUMO for the isolated 1,2-ETC molecule, confirming that they have the same character. This result is in good accordance with our estimate from the energy levels alignment of the isolated systems, *i.e.*, isolated 1,2-ETC molecule (Table 3.7) and ZnO crystal (Figure 3.3), and is very relevant for possible electron-hole dissociation at the molecule-surface interface. This motivated us to investigate the character of the 1,2-ETC- ZnO orbitals around the LUMO-band crossing points as discussed below.

In Figure 4.11-a we show the band structure for the 1,2-ETC- ZnO system along the Γ -Y path

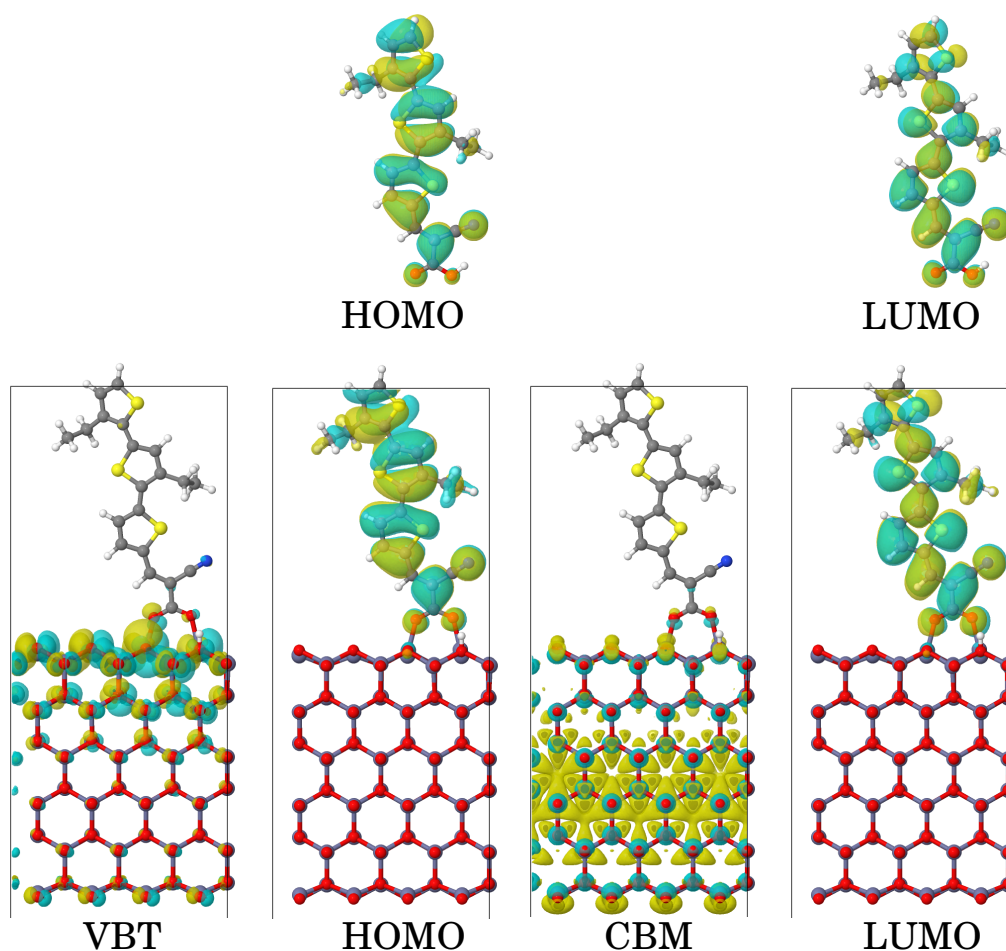


Figure 4.10: Electronic charge densities at the Γ -point for the indicated states of the 1,2-ETC isolated molecule (top) and 1,2-ETC-*ZnO* attached system (bottom) obtained with PBE/ONCVPP. Isosurfaces cutoff of 10^{-4} electrons·bohr⁻³.

(see Figure 4.6-b). The straight black line around 0.8eV concerns the 1,2-ETC-HOMO energy level. It is flat, indicating that the molecule images are not interacting in the molecule-surface model system. Also, it does not interact with the *ZnO* since all states along the Γ -Y path display the same isolated character shown in Figure 4.10. The 1,2-ETC-LUMO energy level is also flat and we show in Figure 4.11-b the corresponding electronic charge densities for five k -points (empty squares) along the Γ -Y path next to Γ . As the LUMO level gets closer to the CBM+2 band (circles), we see a hybridization of the molecular state with the states of the crystal, which is stronger at the LUMO-CBM+2 crossing point. The CBM and CBM+1 bands also bend at Y and cross the LUMO level, and we verify similar hybridization around the LUMO-CBM and LUMO-CBM+1 crossing points. It is worth noting that although the bands for the *ZnO* slab are named CBM+1 and CBM+2, as discussed in the previous section, these bands unfold from the crystal CBM band.

The electronic structure of the 1,2-ETC-*ZnO* system was further investigated using PBE0. In this case, we adopted a wavefunction (electronic density) cutoff energy of 50Ry (200Ry) and,

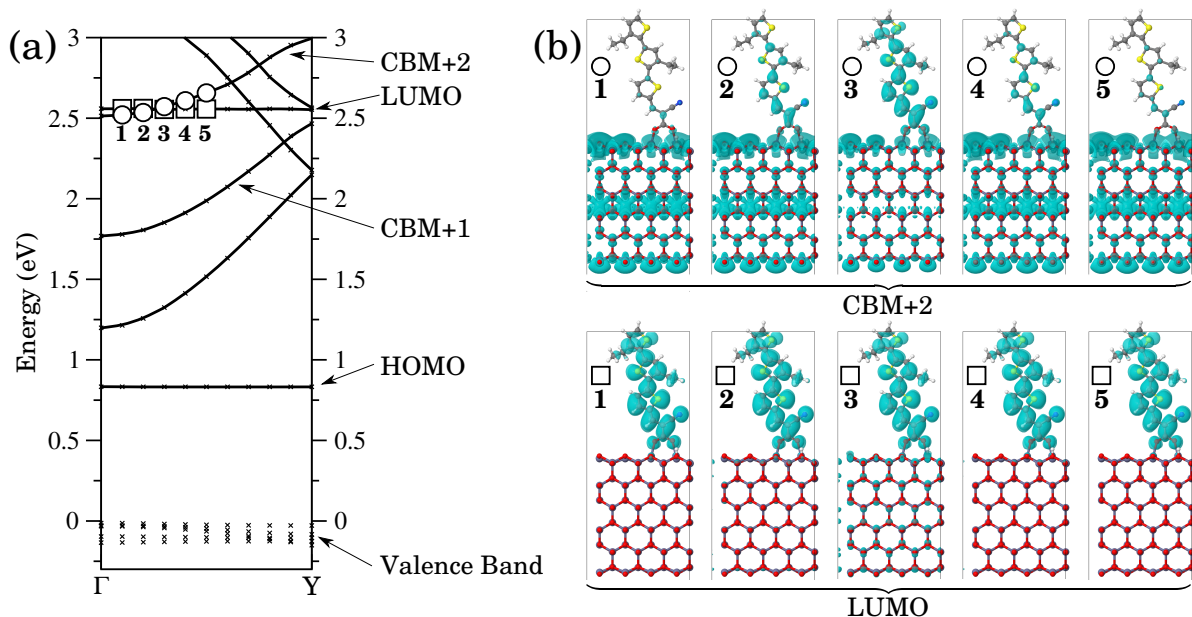


Figure 4.11: (a) Band structure for the 1,2-ETC-*ZnO* system along the Γ -Y path obtained through PBE/ONCVPP calculations. (b) Orbital electronic charge densities concerning the states indicated in “a”. Isosurfaces cutoff of 10^{-4} electrons·bohr $^{-3}$.

concerning the Fock-operator expansion, a cutoff energy of 50Ry. A total-energy-convergence threshold of 1E-3Ry was adopted to converge the SCF cycles. We performed two different calculations: (i) adopting the Γ and Y high-symmetry points (black symbols in Figure 4.12) and (ii) the Γ -point and three other k -points next to Γ (white symbols), placed along the Γ -Y path.

In Figure 4.12-a we present the PBE0 band structure of the 1,2-ETC-*ZnO* system along the Γ -Y path. As we see, the PBE0 results predict an electronic gap of 3.2eV for the *ZnO* surface in good agreement with the reference values [27, 78]. The 1,2-ETC-HOMO-LUMO gap also increases with the inclusion of exact exchange, moving from 1.7eV (PBE) to 3.0eV (PBE0). Regarding the molecule-surface energy levels alignment, we see a similar result in comparison with that for the PBE, namely the HOMO is placed in the *ZnO* gap and the LUMO in the conduction band (CB), near the CBM+2 band. Differently from the PBE results, now two molecular states appear in the *ZnO* gap: the HOMO and HOMO-1. All molecular states are flat indicating that the molecule images are not interacting.

We pass again to the analysis of the states with possible charge transfer, shown in Figure 4.12-b. Exactly at the Γ -point, we see no molecule-crystal electronic communication. However, as we move along the Γ -Y path, we see hybridization of the molecular LUMO with the CBM+2 band around the LUMO-CBM+2 crossing point, similar to the PBE results. The electronic states of the CBM+2 band are particularly interesting since they exhibit electronic charge density spread across the entire slab, favoring charge injection into the crystal. We further show the hybridization of the LUMO with the CBM+1 band at the Y-point. The states of the CBM+1

band exhibit a small electronic charge density concentration in the middle of the slab, indicating reduced molecule-crystal charge injection in this case.

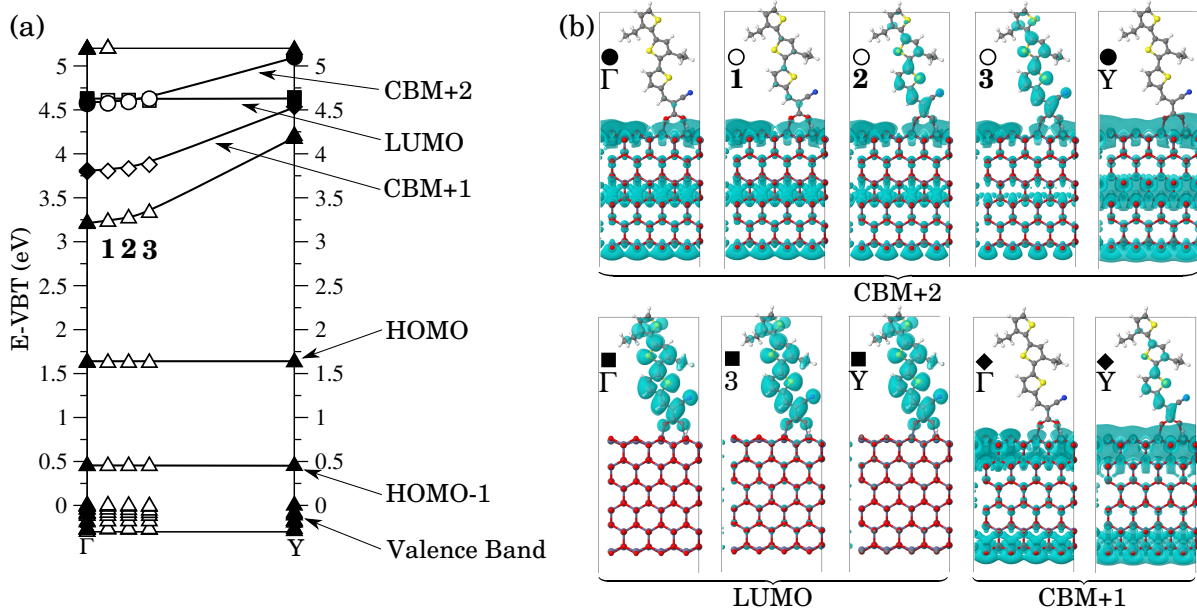


Figure 4.12: (a) Band structure for the 1,2-ETC-ZnO system along the Γ -Y path obtained through PBE0/ONCVPP calculations. VBT set to zero. (b) Orbital electronic charge densities concerning specific states shown in “a”. Isosurfaces cutoff of 10^{-4} electrons·bohr $^{-3}$.

As discussed in Section 3.3, our results obtained using FHI-aims show that the brightest optical excitation for 1,2-ETC happens around 2.5eV (490nm) in the visible range of the solar spectrum. This excitation is due to a HOMO→LUMO transition (95%), where the HOMO is mainly localized on the 1,2-ET moiety and the LUMO on the TC termination (Donor-Acceptor character). This same behavior is observed with QE for the 1,2-ETC-HOMO and LUMO electronic charge densities as shown in Figure 4.13.

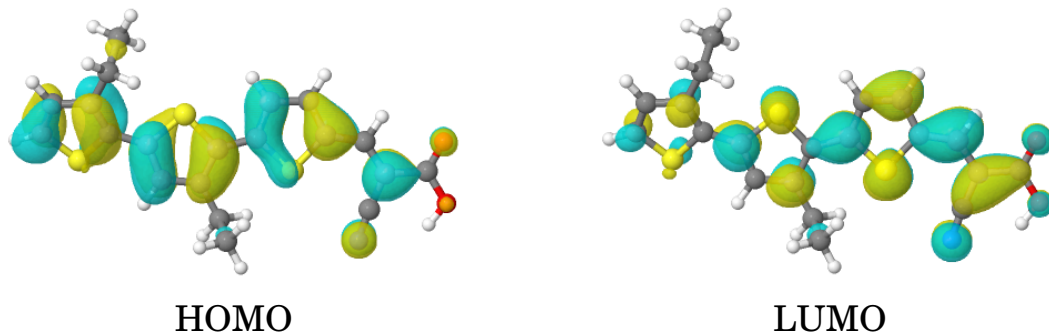


Figure 4.13: 1,2-ETC-HOMO and LUMO electronic charge densities obtained with PBE0/ONCVPP. Isosurfaces cutoff of $3 \cdot 10^{-4}$ electrons·bohr $^{-3}$.

Now, we have investigated the electronic structure properties of the 1,2-ETC molecule an-

chored to the ZnO ($10\bar{1}0$) surface. We find a molecule-surface binding energy of -2.26eV , favorable to chemical bonding. Regarding the molecule-surface energy levels alignment, the 1,2-ETC-HOMO energy level is placed in the ZnO gap and does not communicate with the crystal states, while the 1,2-ETC-LUMO hybridizes with the crystal states in the CB, as shown in Figure 4.14. These states are particularly interesting since they exhibit electronic charge density spread across the entire slab, favoring charge injection into the crystal.

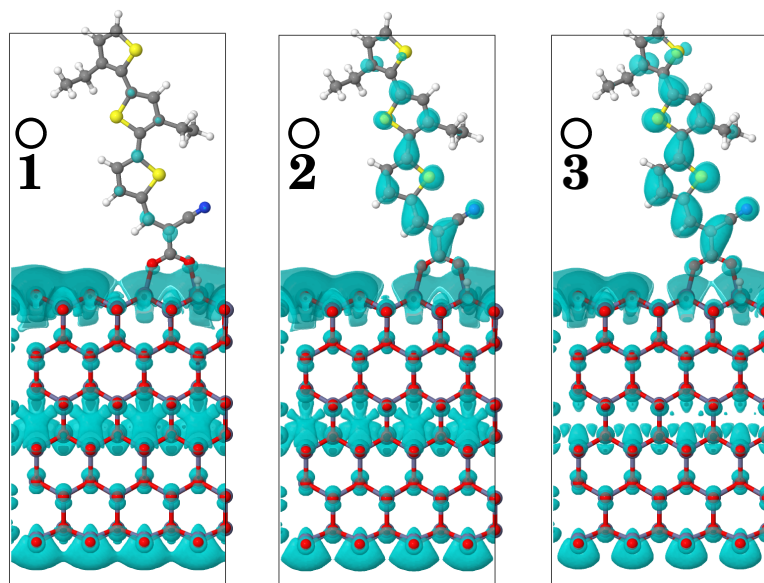


Figure 4.14: Charge-transfer states around the LUMO-CBM+2 crossing point shown in Figure 4.12. Isosurfaces cutoff of 10^{-4} electrons $\cdot\text{bohr}^{-3}$.

Finally, we can expect that once the exciton is formed in the 1,2-ETC molecule through sunlight absorption, it decays into a charge-transfer state with the hole confined on the molecule and the electron extended into the ZnO crystal, with a high probability of electron-hole dissociation. Since we see no communication of the HOMO with the ZnO , electron-hole recombination in the crystal would be prevented. Furthermore, in Section 3.1, we showed that the HOMO energy for oligothiophenes increases with increasing chain size, becoming higher for P3HT than for 1,2-ETC (similar to 3T). Since the hole moves to higher energy levels, we expect favorable hole injection from the 1,2-ETC to P3HT, adopted as the hole-transporting layer in DSSCs. As such, we find that the 1,2-ETC- ZnO architecture shows promising features for application in DSSCs.

Conclusions

In this work, we have performed a theoretical investigation of thiophene-based molecules relevant to the application as an active medium in photovoltaic solar cells. We adopted MP2-corrected HF, DFT, and hybrid HF/DFT methodologies to study the structural and ground-state electronic structure properties for such systems. We have further accessed excited-state properties through the GW perturbation theory and the Bethe-Salpeter equation.

We have investigated two classes of systems: thiophene-furan co-oligomers and cyanoacrylate terminated oligothiophenes. We have shown that thiophene and furan co-polymerization leads to essential modifications in the structural, electronic, and optical properties with respect to thiophene-only oligomers. Regarding the structural properties, while bithiophene stabilizes in a non-planar conformation with an inter-ring torsion angle of 30° or 150° , the thiophene-furan dimer (TF) is more stable in a planar conformation (0° or 180°). This result comes as a consequence of the effective atomic charges interaction between the thiophene and furan rings since the sulfur and oxygen atoms display Hirshfeld charges of opposite sign, namely $0.1e$ and $-0.07e$, respectively, favoring TF planarity. We have further investigated thiophene-furan trimers and tetramers in an antiparallel configuration and show that avoiding TT sequencing in the chain composition leads them to planarity. This prescription can be followed to enhance charge mobility and improve the morphology of thin films made of these compounds.

Thiophene-furan co-oligomers also exhibit interesting electronic and optical properties since they combine the characteristics of the thiophene and furan precursors. Considering chains of the same size, we obtain very similar IP values for the thiophene-furan systems and the furan-only oligomers, whereas, regarding the HOMO-LUMO gap, the results are closer to the thiophene-only oligomers. It is worth noting that thiophene and furan oligomers display similar electronic structure properties, however, the thiophene HOMO-LUMO gap, which dictates the optical gap, is better for sunlight capture. Regarding the optical properties, all thiophene and furan chains (hybrid and non-hybrid) display the first optical excitation as the brightest and mainly HOMO \rightarrow LUMO. These π -states are delocalized over the entire chains and have a similar character for all systems, with electronic charge density localized more on the sulfur than on the oxygen. As commented above, since the thiophene-furan co-oligomers display a HOMO-LUMO gap similar to oligothiophenes, the first optical excitation energy is also closer

to that for the oligothiophenes. The first optical excitation is redshifted with increasing chain size (gap closure) and becomes brighter since the transition dipole moment vector is directed along the chain direction. For the TFTF and TFFT tetramers, the first optical excitation is placed at 3.04eV and 3.13eV, respectively, in the visible range of the solar spectrum. Also, the first optical excitation is brighter for the thiophene-furan co-oligomers compared to 4T. These results highlight the importance of thiophene-furan co-polymerization in the design of new materials for photocurrent generation.

Regarding the cyanoacrylate-terminated oligothiophenes, three systems have been investigated, namely a cyanoacrylated thiophene (TC), and two cyanoacrylated terthiophenes with different alkylation schemes, named here 1,3-OTC (first and third thiophene rings octylated) and 1,2-HTC (first and second thiophene rings hexylated). The TC is more stable in a planar conformation, *i.e.*, torsion angle of 0° or 180° between the T and C moieties. The terthiophene moiety in 1,3-OTC and 1,2-HTC is non-planar but exhibits lower out-of-plane inter-ring torsion angles in 1,2-HTC. These molecules have a Donor-Acceptor character where the thiophene part behaves as the donor and the cyanoacrylate as the acceptor group. While the HOMO electronic density is mainly localized on the 3T moiety in the 3TC systems, the LUMO is mostly on the TC termination. Consequently, the IP for the TC and the 3TC systems is very similar to the respective thiophene-only oligomers, *i.e.*, T and 3T, whereas the EA values are close for all studied *n*TCs and considerably higher than for the oligothiophenes, *e.g.*, around 2.0eV and 0.4eV for the 3TC systems and the 3T, respectively. Accordingly, the oligothiophenes HOMO-LUMO gap closes through cyanoacrylate functionalization, without affecting the IP values. Concerning the optical properties, for all *n*TCs, the first optical excitation is the brightest and it is mainly HOMO \rightarrow LUMO. With the gap closure, this optical excitation is redshifted for all *n*TCs in comparison with that for the oligothiophenes. For the 1,3-ETC and 1,2-ETC (ethylated systems), it is placed at around 2.6eV, in the blue color range, with an exciton binding energy of around 2.6eV, smaller than that found for 3T (3.2eV). The cyanoacrylate functionalization also increases the strength of this optical excitation, whose transition dipole moment vector is directed along the chain direction, as commented above, due to the resultant charge displacement between the HOMO and LUMO states.

All the interesting electronic structure properties displayed by the *n*TCs motivated us to perform DFT and hybrid HF/DFT investigation of the photoexcited charge transfer process in a system composed of a single *n*TC molecule anchored to the *ZnO* (10 $\bar{1}$ 0) surface. Our work was motivated by the experimental results of Oehrlein *et al.* [24] where they reported molecule-crystal photoexcited charge transfer from *n*TCs anchored to *ZnO* nanocrystals, indicating possible application in DSSCs. Here, the molecules were attached to a surface *Zn* site through the $-OH$ group in the cyanoacrylate termination by assuming a dissociation mechanism. To investigate the anchoring, we initially linked the TC molecule which is the termination for any *n*TC system, and the TC-*ZnO* geometry was optimized including the Tkatchenko-Scheffler vdW correction. During optimization, the TC molecule formed a secondary bond with the surface

through the = O radical in the carboxyl group, supporting the proposition of a bidentate binding [25]. We obtained molecule-surface binding energies of -2.29eV for TC and -2.26eV for 1,2-ETC, indicating stable chemical bonding.

The electronic structure for the 1,2-ETC-*ZnO* attached system was investigated using the PBE and PBE0 functionals. All molecular levels are flat, confirming that we are simulating an isolated molecule bonded to the *ZnO* surface. The 1,2-ETC-HOMO energy level appears in the middle of the *ZnO* gap and does not communicate with the surface states, and the 1,2-ETC-LUMO is in the CB and hybridizes with the crystal states (charge-transfer states), in a region of high electron mobility. These charge-transfer states have an electronic charge density that spreads along the molecule (mainly on TC) and across all the *ZnO* slab, favoring deep molecule-crystal charge injection. Since the 1,2-ETC-HOMO does not communicate with the *ZnO*, electron-hole recombination in the crystal is prevented. Also, the HOMO energy is higher for P3HT (hole-transporting layer) than for 1,2-ETC, then favorable hole injection is expected from the 1,2-ETC to P3HT. The 1,2-ETC-*ZnO* architecture thus exhibits suitable properties for application as an active medium in DSSCs.

In summary, we have performed a theoretical study of the structural and optoelectronic properties of short thiophene-furan co-oligomers concerning their application as an active medium in OSCs and cyanoacrylated oligothiophenes as sensitizers for the wurtzite *ZnO* crystal in DSSCs, and we show that they exhibit promising properties for such applications.

Appendix

A.1 Bond lengths and bond angles of 3T and 4T through different methods

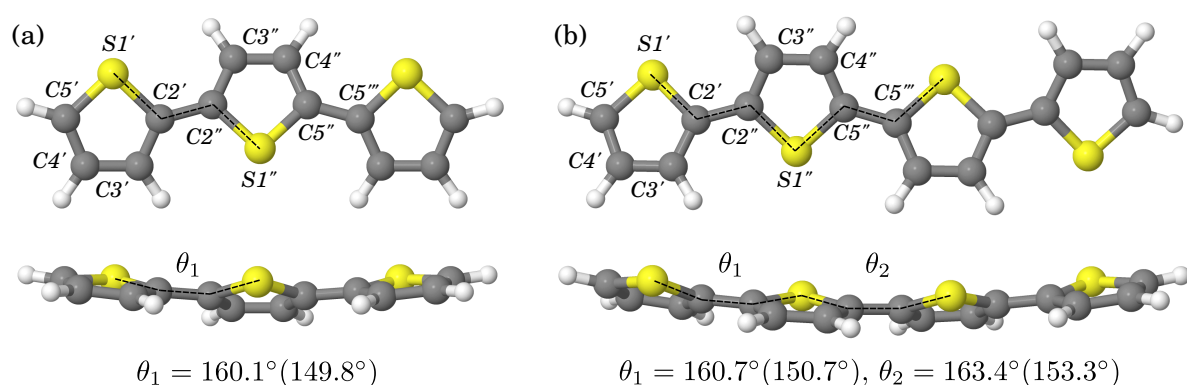


Figure A.1: Top and side views of 3T (a) and 4T (b) geometries obtained using PBE0/6-31g(d,p) with constrained *S-C-C-S* dihedral-angle values at 150° . The presented dihedral-angle values were obtained through further relaxations using PBE0/cc-pVDZ (MP2/cc-pVDZ) with no dihedral constraints.

Table A.1: Bond lengths (Å) and bond angles (°) of 3T and 4T obtained with different methods. The bonds are identified according to Figure A.1.

	3T			4T		
	PBE0	PBE0	MP2	PBE0	PBE0	MP2
	6-31g(d,p)	cc-pVDZ	cc-pVDZ	6-31g(d,p)	cc-pVDZ	cc-pVDZ
$S1' - C2'$	1.74	1.74	1.74	1.74	1.74	1.74
$C2' - C3'$	1.38	1.38	1.39	1.38	1.38	1.39
$C3' - C4'$	1.42	1.42	1.42	1.42	1.42	1.42
$C4' - C5'$	1.37	1.37	1.39	1.37	1.37	1.39
$C5' - S1'$	1.72	1.72	1.72	1.72	1.72	1.72
$C2' - C2''$	1.45	1.45	1.45	1.45	1.45	1.45
$S1' - C2' - C3'$	110.3	110.3	110.6	110.3	110.3	110.6
$C2' - C3' - C4'$	113.2	113.3	113.0	113.2	113.3	113.0
$C3' - C4' - C5'$	112.7	112.7	112.6	112.7	112.7	112.6
$C4' - C5' - S1'$	111.7	111.6	111.7	111.7	111.6	111.7
$C5' - S1' - C2'$	92.1	92.0	92.2	92.0	92.0	92.2
$S1' - C2' - C2''$	120.8	120.7	120.7	120.8	120.7	120.7
$S1'' - C2''$	1.74	1.74	1.74	1.74	1.74	1.74
$C2'' - C3''$	1.38	1.38	1.40	1.38	1.38	1.40
$C3'' - C4''$	1.41	1.41	1.41	1.41	1.41	1.41
$C4'' - C5''$	1.38	1.38	1.40	1.38	1.38	1.40
$C5'' - S1''$	1.74	1.74	1.74	1.74	1.74	1.74
$C5'' - C5'''$	1.45	1.45	1.45	1.44	1.44	1.45
$S1'' - C2'' - C3''$	110.4	110.3	110.6	110.4	110.3	110.6
$C2'' - C3'' - C4''$	113.4	113.5	113.2	113.4	113.5	113.2
$C3'' - C4'' - C5''$	113.4	113.5	113.2	113.4	113.5	113.2
$C4'' - C5'' - S1''$	110.4	110.3	110.6	110.4	110.3	110.5
$C5'' - S1'' - C2''$	92.3	92.4	92.4	92.3	92.3	92.4
$S1'' - C5'' - C5'''$	120.8	120.7	120.7	120.6	120.6	120.6

Bibliography

- [1] H. Hoppe and N. S. Sariciftci. Organic solar cells: an overview. *J. Mater. Res.* , 19(7):1924–1945, 2004.
- [2] D. A. da Silva Filho *et al.* Hole-vibronic coupling in oligothiophenes: impact of backbone torsional flexibility on relaxation energies. *Phil. Trans. Royal Soc.*, 365(1855):1435–1452, 2007.
- [3] O. Gidron *et al.* Towards “green” electronic materials. α -oligofurans as semiconductors. *Chem. Commun.* , 47(7):1976–1978, 2011.
- [4] A. J. Sneyd *et al.* Efficient energy transport in an organic semiconductor mediated by transient exciton delocalization. *Sci. Adv.* , 7(32):eabh4232, 2021.
- [5] A. E. Steen *et al.* Characterization of furan-and thiophene-containing bispyridyl oligomers via spectroscopic, electrochemical, and TD-DFT methods. *J. Phys. Chem. C*, 123(24):15176–15185, 2019.
- [6] O. Gidron *et al.* Study of a bifuran vs. bithiophene unit for the rational design of π -conjugated systems. What have we learned? *Chem. Commun.* , 49(56):6256–6258, 2013.
- [7] G. R. Hutchison, M. A. Ratner, and T. J Marks. Intermolecular charge transfer between heterocyclic oligomers. Effects of heteroatom and molecular packing on hopping transport in organic semiconductors. *J. Am. Chem. Soc.* , 127(48):16866–16881, 2005.
- [8] Y. Qiu *et al.* Synthesis of polyfuran and thiophene-furan alternating copolymers using catalyst-transfer polycondensation. *ACS Macro Lett.* , 5(3):332–336, 2016.
- [9] H. Cao and P. A. Rugar. Recent advances in conjugated furans. *Chem. Eur. J.* , 23(59):14670–14675, 2017.
- [10] O. Gidron, Y. Diskin-Posner, and M. Bendikov. α -oligofurans. *J. Am. Chem. Soc.* , 132(7):2148–2150, 2010. PMID: 20121137.
- [11] O. Gidron and M. Bendikov. α -oligofurans: an emerging class of conjugated oligomers for organic electronics. *Angew. Chem. Int. Ed.* , 53(10):2546–2555, 2014.

- [12] A. Gandini. The irruption of polymers from renewable resources on the scene of macromolecular science and technology. *Green Chem.*, 13(5):1061–1083, 2011.
- [13] H. Tamura *et al.* Theoretical analysis on the optoelectronic properties of single crystals of thiophene-furan-phenylene co-oligomers: efficient photoluminescence due to molecular bending. *J. Phys. Chem. C*, 117(16):8072–8078, 2013.
- [14] B. C. Streifel, J. F. M. Hardigree, H. E. Katz, and J. D. Tovar. Heteroaromatic variation in amorphous 1, 6-methano [10] annulene-based charge-transporting organic semiconductors. *J. Mater. Chem. C*, 2(37):7851–7858, 2014.
- [15] Y. Xiong *et al.* A furan–thiophene-based quinoidal compound: a new class of solution-processable high-performance n-type organic semiconductor. *Advanced Materials*, 28(28):5949–5953, 2016.
- [16] J. C. Sancho-García and A. Karpfen. Theoretical approach to the conformational analysis of heteroaromatic dimers: 2-(2-thienyl) pyrrole, 2-(2-thienyl) furan, and 2-(2-furyl) pyrrole. *Chem. Phys. Lett.*, 473(1-3):49–56, 2009.
- [17] M. A. Perkins, L. M. Cline, and G. S. Tschumper. Torsional profiles of thiophene and furan oligomers: probing the effects of heterogeneity and chain length. *J. Phys. Chem. A*, 125(28):6228–6237, 2021.
- [18] A. Abate *et al.* An organic “donor-free” dye with enhanced open-circuit voltage in solid-state sensitized solar cells. *Adv. Energy Mater.*, 4(13):1400166, 2014.
- [19] M. Planells, A. Abate, H. J. Snaith, and N. Robertson. Oligothiophene interlayer effect on photocurrent generation for hybrid TiO₂/P3HT solar cells. *ACS Appl. Mater. Interfaces*, 6(19):17226–17235, 2014.
- [20] G. H. Lee and Y. S. Kim. Donor-free dyes with an additional acceptor unit for dye-sensitized solar cells. *J. Nanosci. Nanotechnol.*, 17(11):8192–8195, 2017.
- [21] F. Pandolfi, D. Rocco, and L. Mattiello. Synthesis and characterization of new D– π -A and A– π -D– π -A type oligothiophene derivatives. *Org. Biomol. Chem.*, 17(11):3018–3025, 2019.
- [22] J. Li, H. Li, P. Winget, and J. L. Brédas. Electronic structure of the perylene–zinc oxide interface: computational study of photoinduced electron transfer and impact of surface defects. *J. Phys. Chem. C*, 119(33):18843–18858, 2015.
- [23] I. Arbouch *et al.* Influence of the nature of the anchoring group on electron injection processes at dye–titania interfaces. *Phys. Chem. Chem. Phys.*, 19(43):29389–29401, 2017.

- [24] A. N. Oehrlein *et al.* Effect of extending conjugation via thiophene-based oligomers on the excited state electron transfer rates to ZnO nanocrystals. *Phys. Chem. Chem. Phys.* , 21(13):6991–6998, 2019.
- [25] A. Dominguez *et al.* First principles investigations on the electronic structure of anchor groups on ZnO nanowires and surfaces. *J. Appl. Phys.* , 115(20):203720, 2014.
- [26] K. M. Wong *et al.* First-principles investigation of the size-dependent structural stability and electronic properties of O-vacancies at the ZnO polar and non-polar surfaces. *J. Appl. Phys.* , 113(1):014304, 2013.
- [27] A. Janotti and C. G. Van de Walle. Fundamentals of zinc oxide as a semiconductor. *Rep. Prog. Phys.* , 72(12):126501, 2009.
- [28] A. B. Djurišić, X. Liu, and Y. H. Leung. Zinc oxide films and nanomaterials for photovoltaic applications. *Phys. Status Solidi Rapid Res. Lett.*, 8(2):123–132, 2014.
- [29] J. C. Slater. The theory of complex spectra. *Phys. Rev.* , 34(10):1293, 1929.
- [30] H. Bruus and K. Flensberg. *Many-body quantum theory in condensed matter physics: an introduction*. OUP Oxford, 2004.
- [31] J. Toulouse. Introduction to perturbation theory and coupled-cluster theory for electron correlation. https://www.lct.jussieu.fr/pagesperso/toulouse/enseignement/introduction_pt_cc.pdf, 2019. Accessed: 2023-10-09.
- [32] M. L. Krasnov, G. I. Makarenko, A. I. Kiselev, and G. Yankovsky. *Problems and exercises in the calculus of variations*. Mir Publishers Moscow, Russia, 1975.
- [33] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.* , 136(3B):B864, 1964.
- [34] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.* , 140(4A):A1133, 1965.
- [35] T. L. Gilbert. Hohenberg-Kohn theorem for nonlocal external potentials. *Phys. Rev. B*, 12(6):2111, 1975.
- [36] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized Gradient Approximation made simple. *Phys. Rev. Lett.* , 77:3865–3868, 1996.
- [37] C. Adamo and V. Barone. Toward reliable density functional methods without adjustable parameters: The PBE0 model. *J. Chem. Phys.* , 110:6158–6170, 1999.
- [38] C. Møller and M. S. Plesset. Note on an approximation treatment for many-electron systems. *Phys. Rev.* , 46(7):618, 1934.

- [39] M. Valiev *et al.* NWChem: a comprehensive and scalable open-source solution for large scale molecular simulations. *Comput. Phys. Commun.* , 181(9):1477–1489, 2010.
- [40] C. Friedrich and A. Schindlmayr. Many-body perturbation theory: the GW approximation. *NIC series*, 31:335, 2006.
- [41] D. Golze, M. Dvorak, and P. Rinke. The GW compendium: a practical guide to theoretical photoemission spectroscopy. *Front. Chem.*, 7:377, 2019.
- [42] L. Hedin. New method for calculating the one-particle Green’s function with application to the electron-gas problem. *Phys. Rev.* , 139(3A):A796, 1965.
- [43] E. E. Salpeter and H. A. Bethe. A relativistic equation for bound-state problems. *Phys. Rev.* , 84(6):1232, 1951.
- [44] G. Strinati. Effects of dynamical screening on resonances at inner-shell thresholds in semiconductors. *Phys. Rev. B*, 29(10):5718, 1984.
- [45] X. Blase *et al.* The Bethe–Salpeter Equation formalism: from physics to chemistry. *J. Phys. Chem. Lett.*, 11(17):7371–7382, 2020.
- [46] C. Liu *et al.* All-electron ab initio Bethe-Salpeter equation approach to neutral excitations in molecules with numeric atom-centered orbitals. *J. Chem. Phys.* , 152(4):044105, 2020.
- [47] M. Cococcioni and S. de Gironcoli. Linear response approach to the calculation of the effective interaction parameters in the LDA + U method. *Phys. Rev. B*, 71:035105, Jan 2005.
- [48] M. Cococcioni. The LDA+U approach: a simple Hubbard correction for correlated ground states. *Correlated Electrons: From Models to Materials Modeling and Simulation*, 2, 2012.
- [49] R. A. Powell, W. E. Spicer, and J. C. McMenamin. Photoemission studies of wurtzite zinc oxide. *Phys. Rev. B*, 6:3056–3065, Oct 1972.
- [50] H. Li *et al.* Zinc oxide as a model transparent conducting oxide: a theoretical and experimental study of the impact of hydroxylation, vacancies, interstitials, and extrinsic doping on the electronic properties of the polar ZnO (0002) surface. *Chem. Mater.* , 24(15):3044–3055, 2012.
- [51] A. I. Liechtenstein, V. I. Anisimov, and J. Zaanen. Density-functional theory and strong interactions: orbital ordering in Mott-Hubbard insulators. *Phys. Rev. B*, 52(8):R5467, 1995.
- [52] A. Messiah. *Quantum Mechanics Vol. 1*. North Holland Publishing Company, 1964.

- [53] C. Cohen-Tannoudji, B. Diu, and F. Laloë. *Quantum mechanics, volume I*. John Wiley & Sons, 1986.
- [54] A. Szabo and N. S. Ostlund. *Modern quantum chemistry: introduction to advanced electronic structure theory*. Courier Corporation, 2012.
- [55] L. V. C. Assali. Teoria do funcional da densidade. <http://www.fmt.if.usp.br/~lassali/?q=apostilas>, 2020. Accessed: 2023-09-10.
- [56] C. J. Cramer. *Essentials of computational chemistry: theories and models*. John Wiley & Sons, 2013.
- [57] W. J. Hehre, R. Ditchfield, and J. A. Pople. Self-consistent molecular orbital methods. XII. Further extensions of Gaussian-type basis sets for use in molecular orbital studies of organic molecules. *J. Chem. Phys.* , 56(5):2257–2261, 1972.
- [58] T. H. Dunning Jr. Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen. *J. Chem. Phys.* , 90(2):1007–1023, 1989.
- [59] V. Blum *et al.* Ab initio molecular simulations with numeric atom-centered orbitals. *Comput. Phys. Commun.* , 180(11):2175–2196, 2009.
- [60] I. Y. Zhang *et al.* Numeric atom-centered-orbital basis sets with valence-correlation consistency from H to Ar. *New J. Phys.* , 15(12):123033, 2013.
- [61] N. W. Ashcroft and N. D. Mermin. *Solid state physics*. Cengage Learning, 2022.
- [62] H. J. Monkhorst and J. D. Pack. Special points for Brillouin-Zone integrations. *Phys. Rev. B*, 13(12):5188, 1976.
- [63] Occupation methods. https://docs.quantumatk.com/manual/technicalnotes/occupation_methods/occupation_methods.html#occupation-methods. Accessed: 2023-07-31.
- [64] D. R. Hamann, M. Schlüter, and C. Chiang. Norm-conserving pseudopotentials. *Phys. Rev. Lett.* , 43(20):1494, 1979.
- [65] D. Vanderbilt. Soft self-consistent pseudopotentials in a generalized eigenvalue formalism. *Phys. Rev. B*, 41(11):7892, 1990.
- [66] P. E. Blöchl. Projector augmented-wave method. *Phys. Rev. B*, 50(24):17953, 1994.
- [67] D. R. Hamann. Optimized norm-conserving Vanderbilt pseudopotentials. *Phys. Rev. B*, 88(8):085117, 2013.

- [68] J. Yang, L. Z. Tan, and A. M. Rappe. Hybrid functional pseudopotentials. *Phys. Rev. B*, 97(8):085130, 2018.
- [69] Basis set exchange. <https://www.basissetexchange.org/>. Accessed: 2023-07-31.
- [70] NWChem user documentation. <https://nwchemgit.github.io/Home.html>. Accessed: 2023-07-31.
- [71] P. Giannozzi *et al.* Quantum Espresso: a modular and open-source software project for quantum simulations of materials. *J. Phys. Condens. Matter.*, 21(39):395502, 2009.
- [72] P. Giannozzi *et al.* Advanced capabilities for materials modelling with Quantum Espresso. *J. Phys. Condens. Matter.*, 29(46):465901, 2017.
- [73] G. Raos, A. Famulari, and V. Marcon. Computational reinvestigation of the bithiophene torsion potential. *Chem. Phys. Lett.*, 379(3-4):364–372, 2003.
- [74] J. W. G. Bloom and S. E. Wheeler. Benchmark torsional potentials of building blocks for conjugated materials: bifuran, bithiophene, and biselenophene. *J. Chem. Theory Comput.*, 10(9):3647–3655, 2014.
- [75] S. Samdal, E. J. Samuelsen, and H. V. Volden. Molecular conformation of 2,2'-bithiophene determined by gas phase electron diffraction and *ab initio* calculations. *Synth. Met.*, 59(2):259–265, 1993.
- [76] M. Belletete, M. Leclerc, and G. Durocher. Potentialities of semiempirical calculations (AMPAC and INDO/S) in determining the conformation and electronic properties of 2, 2'-bithiophene: a new joint experimental and theoretical approach. *J. Phys. C*, 98(38):9450–9456, 1994.
- [77] G. Liu *et al.* XPS and UPS characterization of the TiO₂/ZnPcGly heterointerface: alignment of energy levels. *J. Phys. Chem. B*, 106(23):5814–5819, 2002.
- [78] S. Blumstengel *et al.* Interface formation and electronic structure of α -sexithiophene on ZnO. *Appl. Phys. Lett.*, 92(19):169, 2008.
- [79] A. Prlj *et al.* Qualitatively incorrect features in the TDDFT spectrum of thiophene-based compounds. *J. Phys. Chem. Lett.*, 6(1):13–21, 2015.
- [80] L. Hung *et al.* Excitation spectra of aromatic molecules within a real-space GW-BSE formalism: role of self-consistency and vertex corrections. *Phys. Rev. B*, 94(8):085125, 2016.
- [81] D. M. P. Holland *et al.* Excited electronic states of thiophene: high resolution photoabsorption Fourier transform spectroscopy and *ab initio* calculations. *Phys. Chem. Chem. Phys.*, 16(39):21629–21644, 2014.

- [82] B. Gombojav *et al.* Temperature dependence of the absorption spectra of bithiophene and terthiophene isolated molecules in cyclodextrin. *J. Phys. Soc. Jpn.* , 73(11):3166–3170, 2004.
- [83] D. V. Lap, D. Grebner, S. Rentsch, and H. Naarmann. Femtosecond-spectroscopic investigations on bithiophene, terthiophene and tetrathiophene in solution. *Chem. Phys. Lett.* , 211(1):135–139, 1993.
- [84] S. Sharma and M. Bendikov. α -oligofurans: a computational study. *Chem. Eur. J.* , 19(39):13127–13139, 2013.
- [85] V. Galasso and N. Trinajstić. A quantum-mechanical study on the stereochemistry of the isomeric bipyrrroles, bifurans, bithiophenes, thienylfurans and biselenophenes. *Tetrahedron Lett.* , 28(16):4419–4429, 1972.
- [86] Jmol/JSmol interactive scripting documentation. <https://chemapps.stolaf.edu/jmol/docs/>. Accessed: 2023-07-31.
- [87] F. L. Hirshfeld. Bonded-atom fragments for describing molecular charge densities. *Theor. Chem. Acta*, 44(2):129–138, 1977.
- [88] L. Klasinc *et al.* Chemistry of excited states. Part 13. Assignment of lowest π -ionizations in photoelectron spectra of thiophen, furan, and pyrrole. *J. Chem. Soc., Perkin Trans. 2*, pages 539–543, 1982.
- [89] D. Klapstein, C. D. MacPherson, and R. T. O’Brien. The photoelectron spectra and electronic structure of 2-carbonyl furans. *Can. J. Chem.*, 68(5):747–754, 1990.
- [90] A. Kraak and H. Wynberg. Charge-transfer interaction of dithienyls and cyclopentadithiophenes with 1, 3, 5-trinitrobenzene (TNB). *Tetrahedron Lett.* , 24(10):3881–3885, 1968.
- [91] I. Novak *et al.* He I photoelectron spectra of bifurans and thienylfurans. *J. Electron Spectrosc. Relat. Phenom.* , 61(1):143 – 147, 1992.
- [92] V. A. Bastos, T. J. da Silva, and M. J. Caldas. Thiophene–furan oligomers: beyond-DFT study of electronic and optical properties. *Electron. Struct.* , 4(1):015004, 2022.
- [93] M. H. Palmer, I. C. Walker, C. C. Ballard, and M. F. Guest. The electronic states of furan studied by VUV absorption, near-threshold electron energy-loss spectroscopy and ab initio multi-reference configuration interaction calculations. *Chem. Phys.* , 192(2):111–125, 1995.
- [94] H. Wynberg, H. J. M. Sinnige, and H. M. J. C. Creemers. Thienylfurans. *J. Org. Chem.* , 36(7):1011–1013, 1971.

- [95] R. Grigg, J. A. Knight, and M. V. Sargent. Studies in furan chemistry. Part IV. 2, 2'-bifurans. *J. Chem. Soc. C*, pages 976–981, 1966.
- [96] A. F. Diaz *et al.* Electrooxidation of aromatic oligomers and conducting polymers. *J. Electroanal. Chem. Interfacial Electrochem.*, 121:355–361, 1981.
- [97] V. Coropceanu *et al.* Charge-transfer electronic states in organic solar cells. *Nat. Rev. Mat.* , 4(11):689–707, 2019.
- [98] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg. A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *J. Chem. Phys.* , 132(15):154104, 2010.
- [99] A. Tkatchenko and M. Scheffler. Accurate molecular van der Waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.* , 102(7):073005, 2009.
- [100] G. J. Martyna and M. E. Tuckerman. A reciprocal space based method for treating long range interactions in ab initio and force-field-based calculations in clusters. *J. Chem. Phys.* , 110(6):2810–2821, 1999.
- [101] S. H. Jin *et al.* Water-soluble thin film transistors and circuits based on amorphous indium–gallium–zinc oxide. *ACS Appl. Mater. Interfaces*, 7(15):8268–8274, 2015.
- [102] A. Dal Corso. Points inside the Brillouin Zone. *Private Communication*.
- [103] H. Morkoç and Ü. Özgür. *Zinc oxide: fundamentals, materials and device technology*. John Wiley & Sons, 2008.
- [104] D. C. Reynolds *et al.* Valence-band ordering in ZnO. *Phys. Rev. B*, 60:2340–2344, Jul 1999.
- [105] M. K. Yaakob *et al.* First principles LDA+U calculations for ZnO materials. *Integr. Ferroelectr.* , 155(1):15–22, 2014.
- [106] O. Dulub, U. Diebold, and G. Kresse. Novel stabilization mechanism on polar surfaces: ZnO (0001)-Zn. *Phys. Rev. Lett.* , 90(1):016102, 2003.
- [107] G. Kresse, O. Dulub, and U. Diebold. Competing stabilization mechanism for the polar ZnO(0001)-Zn surface. *Phys. Rev. B*, 68(24):245409, 2003.
- [108] M. Nilsing, P. Persson, and L. Ojamäe. Anchor group influence on molecule–metal oxide interfaces: periodic hybrid DFT study of pyridine bound to TiO₂ via carboxylic and phosphonic acid. *Chem. Phys. Lett.* , 415(4-6):375–380, 2005.

- [109] R. Lelis-Sousa and M. J. Caldas. Ab initio study of the early stages of gas-phase water oxidation of the Si (100)(2 × 1): H surface. *Phys. Rev. B*, 84(20):205314, 2011.